Collaborative concept mapping and children's learning in primary science.

Sizmur, Stephen Robert

Collaborative Concept Mapping and Children's Learning in Primary Science

Stephen Robert Sizmur

Thesis submitted for the degree of

Doctor of Philosophy

January 1996

King's College, University of London

Abstract

The problem addressed here is that of children's difficulty in learning meaningfully the language of science. Meanings arise in successful acts of communication between individuals, in the context of specific types of activities. Hence communication between pupils using scientific ideas is hypothesized as a worthwhile, effective learning approach.

Concept mapping is suggested as a means of structuring discussion about scientific meanings, and thereby of supporting learning. Although previous research has indicated that concept mapping has a positive effect on pupils' learning, little is known of how this effect is produced or of the contribution made by collaborative group discussion.

This research used qualitative and quantitative approaches to investigate the potential for concept mapping to improve 9- to 11-year-old children's learning in science. It focused on:

- developing appropriate ways of introducing children to concept mapping;
- comparing the effect on learning outcomes from concept mapping with those from a conventional teaching approach;
- comparing the quality of concept maps produced by individuals with those produced by children working collaboratively;
- analysing talk in collaborative concept mapping groups from a sociolinguistic perspective, to identify processes at work in the discussion.

The findings extend those of previous research. They show concept mapping to be beneficial for learning, and to support sustained small-group discussion of scientific ideas. The resulting discourse structure differed substantially from common patterns of classroom talk, often involving children in collaborative construction of relationships shown in the concept maps. When this was the case, the relationships were more likely to be scientifically valid than when primarily the contribution of individuals. Hence collaboratively produced concept maps featured more scientifically appropriate relationships than those constructed individually. However, there was a danger that concept mapping could become decontextualized from other work in science, and suggestions are given for how practice might be improved.

Contents

List of Tables

List of Figures

1

INTRODUCTION

Myopia: is that a place in Yugoslavia?
(Question raised by a Year 5 pupil)

1.1 Overview

This investigation is set in the context of a substantial body of research findings that show children continue to apply scientifically inappropriate ideas in spite of the teaching they experience in school. Unfortunately, not all of these ideas are as easy to detect or as swiftly dealt with as the example above. Rather, they may be deeply ingrained and resistant to change. A burgeoning range of literature in this area testifies to the seriousness with which the problem is regarded (Carmichael *et al.*, 1990; Pfundt & Duit, 1994). Here, then, is the problem to be addressed in the present thesis.

The thesis falls into two main parts. The first is an attempt to elucidate the nature of the problem referred to above, and then to seek, if not solutions, at least measures to alleviate the difficulty. The discussion focuses on what it might mean to learn science in school, and begins with a logical analysis of the nature of science, drawing out certain ideas that are of relevance to learning about science. It then moves on to examine the nature of learning from both a logical and psychological point of view. This results in conceptualizing the problem as one of *lack of meaning* for the learners (Eger, 1992), and there follow implications for the way learning science is approached in the classroom. The second part comprises a practical study of a learning tool known as concept mapping, which, it will be argued, conforms with the ideas of learning presented here.

In this chapter, the educational context of the thesis will be described, and a broad outline of the background theory given. This background theory will be developed further in Chapters 2 and 3, and will underpin the discussion and evaluation of concept mapping as an approach to learning science. In the course of the discussion, so-called "constructivist" approaches to learning (see Driver & Oldham, 1985; von Glasersfeld, 1989) will be subjected to critical appraisal and shown to be inadequate, at least as commonly encountered. It will, however, be shown that what is lacking in this orientation can be made up by drawing on theories that emphasize the shared nature of meaning and the social nature of school learning. This will lead in, Chapter 3, to a reconceptualization of the task of learning science in school, one which locates learning within negotiative interaction between differing viewpoints. As a consequence, it is expected that activities that encourage and structure acts of communication about ideas in science should promote such learning.

In Chapter 4, a learning tool known as concept mapping (Novak & Gowin, 1984) will be described, and its possible suitability for supporting learning of science in the British primary school will be discussed, with reference both to existing empirical evidence and to the theoretical background developed in Chapters 1 to 3. Because concept mapping appears to be a useful way of focusing differing viewpoints about scientific ideas, it is compatible with the conception of learning that underlies this thesis. The discussion in the first four chapters will motivate further empirical research into concept mapping with the following aims:

- to develop appropriate methods for using concept mapping with children in the upper years of National Curriculum Key Stage 2;
- to investigate whether concept mapping executed in this way affects children's learning in science;
- to investigate whether collaboration between children when constructing a concept map contributes to any learning effect identified, and if so to identify and describe the processes involved and how these might contribute to learning.

These aims are elaborated upon in Chapter 5 to create an overall plan and rationale for the empirical study. The responses to the first two of these aims are reported in the latter part of Chapter 5 and in Chapter 6 respectively. The generally positive outcomes warrant further examination of the processes at work to produce the oft-reported learning gains due to concept mapping. Hence, in Chapter 7, we return to the rôle of negotiative interaction in learning with the description of a research design that allows the communicative processes involved in concept mapping to be investigated. The analysis methodology for this main phase of the research, which is set out in Chapter 8, derives directly from the emphasis on communication and meaning.

In Chapter 9, the results of the main phase of the study are presented. In that chapter, specific links are made between the nature of the communication in the groups of children making the concept maps and the understandings that emerge from the activity. In the final chapter, the significance of these findings is discussed in relation to existing research and theory. The conclusions lead to some suggestions for enhancing the use of concept mapping in learning science.

In the next part of the present chapter, some of the main ideas underlying this thesis will be introduced.

1.2 Knowledge: Personal and Public

Knowledge exists first of all in the community. This is socially accepted, public knowledge, which exists independently of any specific persons. But knowledge also exists in the mind of the individual knower, as personal knowledge. Although both types of knowledge may have the same referent, they are not the same. Drawing on the ideas of Wittgenstein, particularly as interpreted and extended by Sainsbury (1992), it will be shown in Chapter 2 that it is only in the context of socially shared meaning that personal meaning can develop. Concept mapping (as described in Chapter 4) operates at the interface of public and personal knowledge, thereby enabling it to play a potentially useful rôle in negotiating new meanings in the classroom.

Here it is useful to make a distinction that will hold throughout this thesis. Those explanatory notions accepted and used within particular cultural groups to classify and understand the world are here termed ***constructs***. These are distinct from the ideas that individual humans (who are members of those groups) hold in their heads. Such ideas are generally termed ***concepts***, and this term will be adopted for the present study. Although the above terminology will be used consistently here, the same words may be used differently by other writers, for whom the distinction is sometimes elided. In fact, the constructs maintained by a society are intimately related to the concepts held by members of the society. The nature of the relationship is what lies at the heart of the present thesis.

1.2.1 Science as Public Knowledge

Scientific knowledge is a particular category of public knowledge. In common with other socially shared meaning systems, science will be portrayed in the following chapter as consisting in sets of practices, each interrelated with a "language-game" (Wittgenstein, 1967). These highly specialized language-games enable new shared meanings to be created and negotiated. Science is above all a *collective* activity. If our task is viewed as bringing pupils to share those meanings, then there are implications for how science must be learned. For to know the meaning of a scientific construct is to know how it is used, and this requires participation in the language-games of science. Hence it will be necessary to distinguish between those characteristics of science that seem to be essential in that they are constitutive of the activity, and those that are peripheral (however significant they may be in practice). One feature of concept mapping is that it reflects some of the characteristics of scientific meaning-making.

1.2.2 Science as Personal Knowledge

Although often compared with the processes of scientific discovery, learning is unlike scientific activity in important ways. Whereas scientific activity is intended to create knowledge that is new to the community, school learning usually is not. Much of the knowledge that teachers intend their pupils to acquire already exists, though it is the product and property of a culture very different to that of the children. Understanding it fully requires an understanding of that culture. Because concept mapping embodies not only characteristics of science as public knowledge but also those of personal knowledge it may have a rôle to play in bridging the gulf between the two.

1.3 The Educational Context of the Study

The National Curriculum for England and Wales requires that pupils of compulsory school age should learn science (GB, DFE, 1995). The outcomes of that learning for children of various ages are specified in the form of attainment targets, and what pupils should be taught to produce those outcomes is also outlined in the form of programmes of study. The Orders for science do not state explicitly what science "is", nor do they embrace any particular theory of how scientific knowledge is acquired by children. But there are certain features of the curriculum (taken here from the programme of study for Key Stage 2) that are worthy of note in developing a view of how science should be approached.

Science is presented as involving systematic enquiry resulting in the acquisition of scientific knowledge, understanding and skills. The scientific ideas that result are to provide explanations for a range of phenomena, and are to be tested against evidence. There are certain scientific constructs specified that children are to learn, accompanied by an increasing grasp of appropriate vocabulary. Separate attainment targets are laid out for the processes of experimental and investigative science and for knowledge and understanding of scientific ideas. From what she refers to as the "packaged, freeze dried curriculum", Harlen (1990) urges educators to produce the "real thing". But what should the "real thing" be like?

Certain value claims underpin this present thesis. One is that children ought to *understand* as fully as possible what they are taught, whether or not this is stated explicitly in the curriculum. Other views of learning are possible, based on different value claims. An alternative philosophy of schooling might eschew understanding, requiring instead verbatim memorization (although such a view would be difficult to sustain under the National Curriculum). The ideas explored in this study would not be of relevance to such a position. Another, related, assumption is that teaching/learning activities are not to be viewed entirely instrumentally, in terms of the results they produce. Empirical studies provide evidence on the *effects* of taking a particular approach; whether, for example, certain kinds of learning outcome are more or less likely. But no empirical study can prove the *value* of a particular learning activity, as what counts as a valuable activity is inextricably tied to the (value-laden) view of learning adopted, and, ultimately, to a view of how persons ought to be treated. This ushers in a third value: that children ought to have their perspective respected, and to be given responsibility for their own learning. The twin aims of this study are to elucidate what ought to take place in learning science, and to apply the view developed to a particular kind of activity; concept mapping.

In the next chapter, an examination will be made, firstly of the nature of the scientific enterprise and of scientific knowledge, and then of some problems and further considerations in developing children's personal understanding of constructs and theories in science.

2

BACKGROUND THEORY

Whereof one cannot speak, thereof one must be silent.
(Ludwig Wittgenstein, 1922, ¶7)

2.1 Scientific Knowledge

In 2.1.1 and 2.2.2, the nature of science as a form of public knowledge will be examined in further detail. It will be argued that certain of the characteristics of scientific knowledge are central to understanding it as a legitimate way of knowing; so central, in fact, that to ignore these in presenting science to children is to risk their rejecting it as incomprehensible and irrelevant. Next, the characteristics of personal knowledge of science will be examined for novices and experts. As novices, pupil's knowledge of science is limited and often inadequate by scientific standards.

The second part of the chapter comprises a critical review of "constructivist" approaches that have been applied in understanding children's difficulties in learning science. It will be shown how some aspects of this perspective can be improved, prior to drawing out implications for enhancing learning.

2.1.1 The Nature of Science

Science may be viewed as a dynamic enterprise, the broad aim of which is to "invent theories that explain observed phenomena" (Kuhn, 1970a, p.2), or "to find theories which, in the light of criticism, get nearer the truth" about the physical universe (Popper, 1970, p.57). Here, "getting nearer the truth" is an important idea, as it dispenses with the popular notion that scientists are in the business of extending the proportion of "the truth" that is known, putting in its stead some kind of ideal limit. However, "truth" is a problematic construct, which will need to be considered with care. There are areas of substantial disagreement between Kuhn and Popper, and also numerous subsequent writers such as Lakatos (1970), regarding precisely how science progresses. Not least amongst these differences is whether there is such a thing as truth in the absolute sense of a single "way the world really is" (a view Putnam, 1981, and others describe as ***metaphysical realism***). There is therefore no agreed view of "what science is". However, common to most is the idea of provisional systems of explanatory constructs, usually termed ***theories***.

Harré (1986) has usefully refocused attention on *that to which scientific theories refer*, which is to say, features of the physical universe. Scientists, he says, do not generally ask "Are the statements of this theory true or false?", but "Do things, properties, processes of this sort exist?" (*ibid.* p.97). Reference does not depend on truth: we can refer successfully with incorrect descriptions. Theories are a means of organising our understanding of the features of the world, and progress within a theoretical position is made when deictic relations are established with some new entity predicted by theory. Ogborn (1995) expresses this well:

> If ... thinking suggests that a certain being, perhaps a gene, an active site on an enzyme, or dislocations in solids, may usefully enter an account of the world, then it is a rational way of going on to mount a search for such a thing. Such an action may succeed or fail. This success or failure is where we draw the line ... between what we like to think and what we can or cannot do. (p.7)

Harré's is a realist view of science. It is dependent, *inter alia,* on the existence of a real physical universe that is relatively more stable than our understanding of it, and which therefore affords the possibility of consistent reference. But it is also a view in which theory is constructive and progressive, suggesting experiences we have not yet had. Moreover, scientific theory produces constructs, *beyond* possible direct experience, that are needed to preserve coherence. Such, for example, are the notion of "acceleration" when predicated of a body at rest (see Eger, 1992) and ideas such as "niche" in biology. *Theories, therefore, cannot be derived directly from experience.* Hence Kuhn (1970c) observed that theories are always underdetermined by data, and involve an arbitrary, creative element. The power of a new theory is that it enables us to break free from the restrictions imposed by existing ways of seeing the world, which Kuhn likens to a Gestalt shift. Both he and Popper (*op cit.*) bear witness to the tremendous intellectual achievement that this entails.

The arbitrary element to scientific theory means that it is not helpful to dwell on whether or not a theory is "true". For one thing, we could never *know* whether or not the "absolute truth" about anything had been discovered, since verification implies a viewpoint external to, and independent of, any theory system (what Putnam, echoing Kant, 1993, calls the "God's Eye" point of view). Popper's project was to insert falsification in place of the unattainable quest for verification. However, as Lakatos (1970) has pointed out, the observations that we depend upon for falsifying evidence are themselves theory-dependent and fallible. As a consequence, we can choose to doubt an observation, and offer hypotheses to explain it away. All theories, he claims, produce anomalies, and an anomaly is not in itself sufficient to overthrow a theory. What makes an experiment crucial is its capacity to discriminate between competing theories, not "falsification" alone. Lakatos therefore proposes that to adjudicate between theories we ask how *successful* a theory is, and whether it is *progressive*. He provides an illustration.

> Einstein's theory is not better than Newton's *because* Newton's theory was "refuted" but Einstein's was not: there are many known "anomalies" to Einsteinian theory. Einstein's theory is better than - that is, represents progress compared with - Newton's theory ... *because* it explained everything that Newton's theory had successfully explained, and it explained also *to some extent* some known anomalies and, in addition, forbade events ... about which Newton's theory had said nothing but which had been permitted by other well-corroborated scientific theories of the day ... (p.124)

From the above description, Einstein's theory is clearly not "true" in any absolute sense. But it is rationally acceptable, and we can therefore grant it "epistemic approval" (Harré, 1986). The everyday notion of "truth" may continue to function as an evaluative principle and as a moral imperative, so long as we abandon the idea that we can assign, once-and-for-all, the property "true" to the statements of a theory. Putnam (*op cit.*) proposes four criteria (four values) which we apply to decide on the epistemic acceptability of theories: ***comprehensiveness***; ***functional simplicity***; ***instrumental efficacy***; and ***coherence***. In essence, these mean that a good theory:

- can account for a relatively wide range of phenomena;
- does so in as economical and elegant a way as possible, without resort to *ad hoc* elements;
- is consistent with available evidence, resulting in accurate predictions (accepting that anomalies will arise from time to time);
- is logically consistent, containing no mutually contradictory propositions.

It is against criteria such as these that rational theory adjudication can take place and scientific progress can be recognized.

It only really makes sense to speak of truth from *within* a theoretical position. For Wittgenstein (1967), truth is established with respect to criteria, or sets of procedures, that are internal to a "language-game". A language-game is his term for "language and the actions into which it is woven" (p.5). Within such a language-game, the meaning of a term is made determinate through its *relationship* to certain other terms and to practices. Thus, if we say something is entirely "red", we know that it is not at the same time blue or green, but we do not know, for example, whether it is large or small, light or heavy (Harrison, 1979). And so "the language game gives us a criterion for dividing sensory events into those that are relevant to the truth of [X], those that are relevant to the truth of [NOT X], and those that are irrelevant to 'the issue of its truth or falsity'" (*ibid.*, p.248).

This simple example, though, does not capture the full significance of the point that truth conditions are internal to a language-game. Winograd (1985) provides a more subtle example:

> Consider the following dialogue:
> A: I'm thirsty
> B: There's some water in the refrigerator
> A: Where? I don't see it
> B: In the cells of the eggplant (p.186)

Speaking from the viewpoint of scientific theory, what B says is perfectly true. Yet we cannot help feeling that B has misled or tricked A, that B is not, in fact, telling the truth. This is because the social context in which the words are uttered (the language-game) requires that B help A overcome a particular need (thirst) and that therefore any response on B's part should be interpreted accordingly. Moreover, one could foresee a situation in which B directs A instead to a liquid that is not pure water (tap water, for example), yet which B refers to quite appropriately as "water" without further qualification. It is not difficult to appreciate why a demand for "absolute", metaphysical realist, truth in all circumstances (assuming such a thing to be attainable) would render much of our everyday activity unworkable. So the "internalist" notion of truth has an important consequence. It allows for the possibility of alternative true, but incompatible, descriptions. Yet it will not yield just any descriptions, as Wittgenstein notes:

> "So you are saying that human agreement decides what is true and what is false?" - It is what human beings *say* that is true and false; and they agree in the *language* they use. That is not agreement in opinions but in form of life (*op cit.*, ¶241).

The force of what Wittgenstein is saying here is that it is through the *actions we perform* in relation to the (obdurate) world that intersubjective judgements confirm or deny the validity of our words.

What lessons can we draw from this brief discussion that will be of relevance to learning science? Firstly, that it is in the interrelationship between language, actions and the unyielding physical world that scientific theories are created. There must be *evidence* for what we say in a theory. Secondly, that the empirical data do not uniquely determine the theory, which has a creative element. How a theory derives from the data is not obvious, and given the same data, we will not all discover the same theory. Thirdly, the development of better theories leads to the discovery of new knowledge of things, properties and processes. Lastly, we should not assume that science is unique in allowing us to tell the truth about the world. There are everyday ways of talking truthfully about the world that conflict with scientific interpretations.

2.1.2 Science as a Social Activity

Kuhn (1970b) has drawn attention to the importance of viewing science as the product of particular communities of individuals. For him, "the very idea of scientific knowledge as a private product presents the same intrinsic problems as the notion of a private language" (*ibid.*, p.253). It is worth exploring what such a statement implies.

Wittgenstein (1967) dismissed the possibility of a private language by arguing that the meaning of any word must conform to public criteria. There could be no communication without agreement on meaning, and, as the quotation in the preceding section shows, this agreement is achieved through joint actions. *The meaning of the word is how it is used,* but use (as was pointed out above) must be considered within a socially constituted "language-game". When we use language, we do so in the context of certain social practices whilst drawing on the resources of the whole language community, including its theory system. Against this public theory system, what we say is open to correction. In science, it is through the community of scientists that the criteria of correctness are applied and theory evaluation takes place. Thus, Knorr-Cetina (1981) says that "the communicative foundation of science constitutes the scientists' operations as a form of *discursive interaction* directed at and sustained by the arguments of others" (p.14). If scientific theories have indeed progressed then this has in large measure been through the critical exchange of ideas within the scientific community, which acts in the manner of "natural selection" (*ibid.*). Acceptance by the community is just as essential as empirical success.

Science is a form of life with a distinctive set of language-games, intimately related to its activities and practices, and through which its ideas must be formulated, analysed and refined. Access to these language-games is neither automatic nor effortless; it is a part of what becoming a scientist involves. And if there is no possibility of stepping outside the theory system to view from a distance its rules of usage (that is, there is no possibility of "absolute", externalist truth), then it is clear that to learn how to play the language-game, one must work *within* it.

In this section, the scientific enterprise has been portrayed as the development of a *shared* system of coherent, but provisional, constructs that explain how the universe functions. Its potency lies in its capacity to subsume more and more of our real experience under these powerful constructs and principles. So one aim of a "real" curriculum in science should be to initiate pupils into that shared understanding, but without denying them other "ways of worldmaking" (Goodman, 1978). In other words, the aim is to help individual pupils to enter the language-games of science. But to enter a language-game is to take up, not its words alone, but its ways of using words in relation to its activities, its ways of seeing the world. This, the subject of Chapter 3, will entail working with key scientific constructs and coming to understand their interrelationships and use within the theory system. However, scientific theories, because they are powerful, are also abstract and, it seems, difficult for many pupils to grasp. This is the subject of the next section.

2.1.3 Children's Understanding of Science

There is now a body of evidence suggesting that the scientific understanding of many children fails to develop beyond a relatively naïve level. The evidence has accrued from what may conveniently be termed the "alternative conceptions movement" (Millar, 1989). This is a research programme in science education, also known as "children's science" (Osborne & Freyberg, 1985) or "alternative frameworks" research (Driver and Easley, 1978).

The central conclusion of this research programme is that children begin school with firmly held beliefs about certain science topics that differ in significant respects from those accepted by the scientific community. They subsequently draw upon these beliefs to make sense of new experiences, including experiences of science teaching in schools. Furthermore, these ideas are not simply random mistakes, but are the result of genuine attempts to interpret new events. Like scientific theories, they are rarely challenged, and are resistant to attempts to dislodge them. (Driver, 1983; Osborne & Wittrock, 1983; Eylon & Linn, 1988).

It has been suggested that children's alternative conceptions form part of a personal conceptual structure or theory system that is capable of providing them with coherent explanations of the world (Osborne & Wittrock, *op cit.*). However, the degree to which these ideas are coherent is much more restricted than is the case for a scientific theory (Driver, Guesne & Tiberghien, 1985; Snir, 1991). Children typically hold multiple conceptions for phenomena that may be explained by a single accepted scientific construct, and the one selected will depend on the context in which a phenomenon is encountered. Some of these alternatives may even be mutually contradictory (Driver *et al.*, *op cit.*). Others of children's conceptions are undifferentiated, and may embrace aspects of a number of distinct scientific ideas (*ibid.*). In these cases, a term may take on different meanings in different contexts. Children's views of the world, then, frequently measure up poorly against Putnam's (1981) criteria of comprehensiveness, functional simplicity, instrumental efficacy and coherence. Within the alternative conceptions movement, the task for the science educator is therefore typically viewed as that of leading children to reject their restricted concepts and to replace them with more consistent, coherent, broader and reliable concepts. This is a view that will be called into question below.

For the present, it is necessary to observe that children's alternative conceptions make the results of teaching unpredictable. Learners may seem to acquire "correct" concepts, yet in unfamiliar situations resort to their intuitive versions. Sometimes it is even possible to detect a regression in children's ideas over time, vis-à-vis accepted scientific views (Osborne & Wittrock, *op cit.*).

As discussed above, scientific constructs are the products and property of a community. In order to help children become knowledgeable about science, it is necessary to consider how an individual may personally appropriate those constructs, retain them and apply them to their experience. Central to what follows is the idea that theories are organized networks of constructs used to categorize experience, and that the way those networks are organized is important in how they are learned, retained and used.

2.1.4 Expert Knowledge

Another substantial field of research has been centred on the nature of expert knowledge, and how it differs from that of novices (for reviews see Eylon & Linn, 1988; Rowell & Dawson, 1989). In this section, we will consider just one important finding from this body of work.

Experts, it seems, organize their knowledge hierarchically, with more general concepts and principles subsuming knowledge of specific instances (Rowell & Dawson, *op cit.*). Eylon & Reif (1984) examined the effects of teaching explicitly the hierarchical structure of knowledge required to solve a class of physics problems. They found that students taught in this way performed substantially better on related problem-solving tasks than students who had been taught more conventionally. Hierarchically organized knowledge apparently enables experts to categorize problems by more relevant, but less salient, cues than novices, who tend to base their categorizations on surface features (Eylon & Linn, *op cit.*; Smith, 1992). It also seems that knowledge that is more cohesively interrelated is easier to retain and recall (Gagné & White, 1978; Chi & Koeske, 1983).

Smith (*op cit.*), however, also found that there was not necessarily a single organization that was optimal for all the different kinds of expertise possible within a domain (compare here; different language-games). Experts in a domain (in this case, genetics) tended to organize knowledge differently according to how it was most frequently called upon. The conclusion drawn by Smith (*ibid.*) is that pupils are best served by equipping them, not with any specific knowledge structure, but with the capability to structure their knowledge to meet their needs.

2.2 Understanding Personal Knowledge

How, then, can we give children the opportunity to acquire appropriately organized, coherent sets of ideas about scientific theories? This question concerns the relationship between public and personal knowledge. In this part of the chapter, a theoretical foundation will be prepared, on which an approach to the teaching and learning of conceptual understanding in science may be based. In doing so, a critical examination will be made of theoretical positions that have previously been advanced as a basis for thinking about children's learning. It will then be shown how reconceptualizing the task might alleviate some of the difficulties this examination raises.

2.2.1 Constructivism

The epistemological and psychological bases of the "alternative conceptions movement", and also of various attempts to find ways to improve children's learning, lie in the position known as ***constructivism.*** Constructivism, though, has not found universal favour, and has come increasingly under attack through the early 1990s (Suchting, 1992; Matthews, 1992; Osborne, 1993). It is therefore apposite to evaluate the views that have been put forward under this banner.

Constructivism is not a unitary theory (Wheatley, 1991), which can make it difficult to pin down. One key notion seems to be that humans "make sense" of events in terms of their prior knowledge. Here, an influential constructivist, von Glasersfeld (1989), expresses the view that:

> the interpretation of experience and language ... have one important feature in common. Both rely on the use of conceptual material that the interpreter must already have. "Making sense", in both cases, means finding a way of fitting

> available conceptual elements into a pattern that is circumscribed by specific constraints. (p.11)

Another common thread is the view that "we cannot contact an interpretation-free reality directly" (Bannister & Fransella, 1971, p.18), which means that we cannot take the "God's eye" viewpoint and see whether the interpretation we construct matches independent reality. Suchting (*op cit.*) attempts to clarify what is meant by constructivism, suggesting that:

- it denies the possibility of knowledge that corresponds with the world-in-itself;
- it claims that knowable reality is the experience of an individual subject, and constructed by that subject;
- it claims that knowledge consists of concepts, and is a mapping of what is viable in experience.

This is a view that runs into serious difficulties, once we begin to ask how it is possible to construct knowledge out of subjective experience. Von Glasersfeld reveals empiricist leanings:

> the compound of experiential elements that constitutes the concept an individual has associated with a word cannot be anything but a compound of abstractions from that individual's own experience. For each of us, then, the meaning of the word apple is an abstraction that he or she made individually from whatever apple experiences he or she has had in the past. That is to say, it is subjective in origin and resides in the subject's head, not in the word that ... has the power to call up, in each of us, our own subjective representation. (p.9)

This is a highly individualistic (one might say solipsistic) and somewhat passive view of knowledge construction. Whilst not denying the existence of an external reality, it draws a sharp line between that reality and the knower. Contrast the preceding with the view of "basic realism" as summarised by Lakoff (1987), which includes:

- commitment to a real world independent of human beings, but including the reality of human experience;
- a link between human conceptual systems and other aspects of reality;
- commitment to stable knowledge of the world;
- rejection of the view that all conceptual systems are equally good.

These include some of the very things denied by von Glasersfeld. By separating out the personal from the social, the active from the passive aspects of knowledge construction, it is possible to avoid the fallacious argument in some constructivist writing, identified by Matthews (*op cit.*), that proceeds from premises about the constructive nature of knowledge acquisition to the conclusion that we cannot know reality. Before proceeding with this, though, attention will be turned to a psychological model of learning, based on constructivist principles, that has proved popular in the "alternative conceptions" movement.

2.2.2 A Constructivist Model of Learning

The Generative Learning Model, as proposed by Osborne & Wittrock, (1983), is a model of human information processing which is purported to account for the manner in which information held in long term memory influences a person's understanding of a situation and subsequent learning.

According to Osborne & Wittrock (*op cit.*), to understand a situation (such as a science lesson) is to generate an adequate internal "model" or explanation of it. It is this that comprises the "meaning" of the experience for the individual, and to which he or she responds. However, the authors do not elaborate on its form.

On encountering a situation, the mass of sensory stimuli received by the individual is not processed simultaneously. According to the model, information stored in long term memory largely determines which stimuli are deemed worthy of attention. Where a stimulus can be connected with information stored in memory, then the stimulus is recognized or ***perceived*** and brought into a (limited capacity) short term store for processing. Other information that appears relevant is also retrieved, and used to generate a tentative model of the situation. This initial construction of meaning is evaluated. If it is viable, if it matches both current and past experience adequately, information about the situation is transferred into long term storage. Learning is then said to have occurred. If the first attempt at generating meaning fails, if it does not match some aspect of past or present experience, then alternative ways of making the connection are explored. This could be achieved by attending to different aspects of the situation, or by attempting to create links with different elements in memory. If seeking alternatives does not seem worthwhile, then the attempt to assign meaning may be abandoned altogether.

The Generative Learning Model enlarges upon how humans attempt to understand reality. It makes use of the important idea that *observations of reality do not contain all the information necessary for an explanation to be constructed, and an interpretation to be made.* The internal model must draw not only on sensory impressions, but also on relevant information in memory. It is this additional information that helps to "fill in the gaps" in the internal model and enables predictions to be made that form the basis of a response.

2.2.3 Limitations of the Model

Although the Generative Learning Model goes some way towards describing how information in memory influences the interpretation of new situations, it leaves a number of other questions unanswered. There is, for example, nothing to indicate how it is that mutually contradictory ideas can coexist, without this being seen by the learner as problematic. Most seriously, the model also seems to assume that our meanings are essentially private, and derived from personal experience, a viewpoint apparently supported by von Glasersfeld (1989) and Wheatley (1991), but shown to be incoherent by Wittgenstein's "private language" argument referred to in 2.1.2. It is the purpose of this and the following sections to address some of these unanswered problems.

A significant difficulty with the article by Osborne & Wittrock (*op cit.*) is that the picture of their psychological model is painted in broad brush strokes. Hence, although the overall impression may be convincing, nevertheless, in many respects it is vague. The idea of "experience" is unanalysed, and appears to refer not only to "an event in which we are involved", but also to something like "raw sense data", which is how von Glasersfeld (*op cit.*) seems to use the term. But experience is intentional, linked to something beyond ourselves:

> human experience is meaningful experience. Our experience consists of interrelated perceptions, thoughts, communications and actions. All this experience is experience *of* something and that something is characterised in a particular way *before* we can be said to experience it at all. (Sainsbury, 1992, p.9; emphasis changed)

Knowledge, and language, cannot be grounded in sense impressions, but are constructed around shared engagement with the (obdurate) world.

Possibly as a result of the failure to analyse "experience", the twin processes of selective attention and perception are not well explained in the Generative Learning Model. According to the authors, construction of meaning begins, not with the experience, but with selective attention to that experience, which " ... requires voluntarily controlled effort" (Osborne & Wittrock, *op cit.*, p.494). But if we only perceive what we attend to, how do we know what to attend to? In order for expectations to be generated and attention to be directed, there must first be some appreciation of the situation (involving an application of meaning) on which to base them. There seems, therefore, to be a circular impasse in this description.

What we attend to are not, therefore, the meaningless sensations assumed within the model, but those aspects of already meaningful experience that seem to require (and to be worthy of) further processing. Where the authors are correct is in concluding that our existing knowledge actively contributes to our interpretation of any given situation, filling in with assumptions (perhaps wrong assumptions) wherever we lack explicit information from our surroundings, and checking out those assumptions in the light of further evidence. Where they are wrong is in following countless other psychological theories that place meaning making primarily within the individual. The metaphor of the child as a lone scientist, observing, hypothesizing, experimenting and concluding, is central, and the rôle of social interaction and the communication of ideas peripheral. Yet, as has been pointed out above, no scientist functions in that way. Science is *necessarily* social, and draws on a *shared* theory system already in existence to formulate and to check its hypotheses. Even if moving into a new paradigm, scientists must make some use of existing shared meanings, without which their ideas could never be communicated. It was Wittgenstein, developing the position now known as "social constructionism" (Lemke, 1990), who reversed the centuries old line of thinking that nevertheless still underlies much of psychology. Reference has already been made to his denial of the possibility of a private language. It is now apposite to develop this approach further.

2.2.4 The Necessity of a Shared Meaning System

The philosopher Marian Sainsbury has explored in detail the social nature of meaning, and what follows in this section draws on her analysis (Sainsbury, 1992).

Meaning may be considered as "a set of classifications" (*ibid.*, p.15). But in making a classification, one must do so with reference to a rule. Objects that are similar in one respect will differ in others, so, as Rorty (1980) points out, we cannot form a concept of something just because we have noticed that kind of thing, since this implies "notice under a description". A rule of some sort is needed to specify what it is for something to be a member of a particular class, that is, what counts as relevant evidence. Such sets of rules were referred to above as intrinsic to "language-games" (2.1.1).

If we wish to make classifications, this entails some notion of correctness. This notion can only be realized with reference to shared agreement over classifications, otherwise we could not know if we were being consistent. Without such consistency, we could never communicate about anything or mean anything. On the other hand, when we learn a language, we are, *ipso facto*, learning to operate the system of rules it embodies: "meaning is not imposed upon the words by the user; it is imposed upon the user by the words" (Sainsbury, *op cit.*, p.21).

These rules are interconnected in a complex set of relationships which give words their meaning. "The search for determinate meaning comes in the end to connections and relationships, and those relationships to a complex set of intersubjective agreements" (*ibid.*, p.46). Constructs (as the term is used in this thesis) are the nodes in this network of relationships. The meaning of a construct resides only in the way it is related to others. This is most important, as it implies that we cannot conceive of constructs in isolation.

This abbreviated account of the logically prior nature of socially shared meaning has a considerable bearing on the constructivist model of learning outlined above. In the first place, it has much to say about the nature of a person's preconceptions. Such conceptions are not logically private, but draw upon the pool of shared meaning in which the individual is immersed from birth. If this were not so, that individual could not function as a member of a culture. Extending the idea that the meaning of a word is its use in a language-game, it is possible to envisage that the same word might occupy a position in a number of language-games, and hence derive its different senses from different sets of relationships. If such language-games were essentially separate activities, then it would be possible for the different senses of the word to have incompatible meanings. There will appear no conflict, provided that they have different contexts of use. (And even when these contexts overlap, there is, as has been shown above, the possibility of several true, but incommensurable, descriptions.)

2.2.5 The Origins of Alternative Conceptions

Each person participates in a wide range of different activities. In the context of these activities, different language-games operate, and one purpose of these language-games is to facilitate, or even constitute, the activities. To the extent that the underlying theory system is coherent, supportive of the activities and in accord with experience, there is normally no need to question its truth.

Hence we can speak meaningfully of "letting the cold in", of vacuum cleaners' "sucking up" dust and of people behaving "like animals" in our everyday encounters. Yet in scientific discourse, these descriptions would not count as being adequate. Our use of certain words can differ in different kinds of discourse, but each time their use derives from a public meaning system. It is when we try to extend a particular sense of a word to a new scenario that we can run into trouble. In science, words can take on new and very different senses. This point echoes that made by Solomon (1983) that pupils inhabit a "life world" and a "symbolic" domain, though in this present account, the emphasis is on the possibility of a multiplicity of different contexts of use.

Some writers have remarked that children's alternative conceptions can resemble in some (limited) respects those that were once held as true by the scientific community (Driver, *et al.*, 1985). Where there is a resemblance, this may also be traceable to public meaning systems, for language is fundamentally metaphorical, and "the literal consists in forgotten metaphor" (Weinsheimer, 1985, p.239). Sutton (1992) describes the interaction between the everyday use of words and their use in scientific discourse. Where meanings from an earlier scientific paradigm persist in everyday linguistic metaphors, elements of the older way of thinking may similarly persist, such as the idea of an "inner light" emanating through the eyes which is suggested by expressions like "his eyes lit up" and "her eyes burned right into me". Further, there is evidence that where languages differ in the metaphors they employ, this results in different sets of alternative conceptions. For example, Hewson (1985) reports that in the

African Sotho culture, people are described as "hot" when they are agitated. This culture does not appear to have developed the notion of heat as a fluid held by many Europeans. On the other hand, the Sotho appear to accept kinetic models quite readily.

Given their location in successful day to day communication, it should be no surprise that alternative conceptions are resistant to change. The main thrust of approaches that seek to displace alternative conceptions is to create conflict and dissatisfaction with them, but it is by no means clear how meaning systems that generally function adequately can be dislodged by isolated cases of discrepant evidence. Indeed, Solomon (*op cit.*) insists that it does children a disservice to try to obliterate such ideas, since this would deprive them of the ability to communicate in everyday life.

This has been a brief account of some of the origins of our context-specific meanings. It does not do justice to the topic. No mention has been made here of the many other factors which probably play a part in our construction of meaning; our natural bias towards confirming our existing theories (Kuhn, Amsel & O'Loughlin, 1988); focusing on change rather than equilibrium (Driver *et al.*, 1985). It is not the purpose of this thesis to deny that such factors are influential, but rather to point out that alternative conceptions are not simply the result of individual experimentation on the world. Consequently, it is not solely to individual experimentation that we must turn to free children from the limitations they impose.

2.2.6 Cognitive Structure

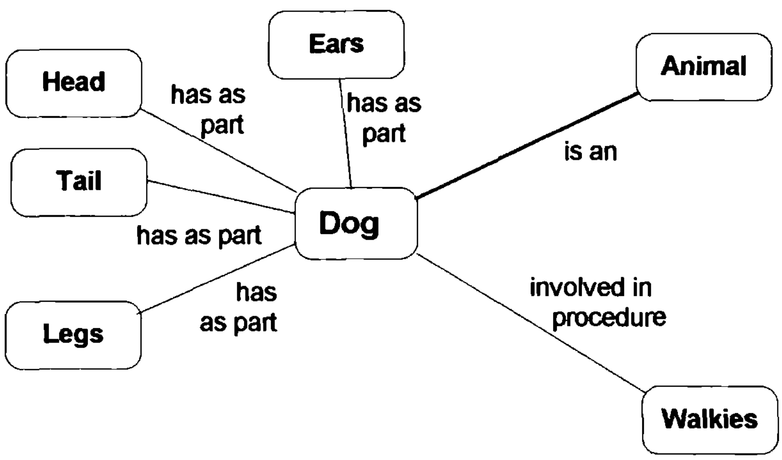
In Sainsbury's account, personal meaning is characterized as "the individual participation in the shared meaning system" (Sainsbury, 1992, p.81). This entails that humans must have some means of "holding on to" aspects of the public meaning system. This is the rôle of long term memory.

Modern theories about memory portray it as having a structure, but this may have one of two senses. The first is the notion that elements in memory are linked, so that during recall activation of one element in the brain causes activation of other associated elements (cf. Collins & Quillian, 1969). Presumably some mechanism of this sort exists, but it is not clear how this helps us to ascribe *meaning* to a situation. The other kind of structure posited is therefore the set of connections between concepts that determines how they relate to each other in terms of use, and consequently in terms of meaning. Typically, models of this kind have knowledge in memory structured in terms of ***propositions***. White (1988), for example, defines propositions as representations of meaningful relationships or properties. Although expressed in symbols, it is the relationship itself (the rule), rather than its expression, that constitutes the memory element. Two propositions may be connected by means of a concept common to both. For example, the proposition "snow is frozen water" is connected to the proposition "snow is white" in virtue of their both being about the concept "snow". The concept itself, though, has no separate existence. It is a point of intersection in a web of meaningful relationships.

This latter is quite a different level of structure to that of quasi-neural pathways or "associations" involved in calling such ideas to mind, and is of central concern in this discussion. Cognitive structure in this sense is the stuff of personal knowledge, and is the internal correlate of the external theory system (public knowledge) that has been invoked in previous sections, where "the search for determinate meaning comes in the end to connections and relationships" (Sainsbury, *op cit.*, p.46).

Consider the propositional knowledge that a typical person might have regarding the class of objects called "dog". Let Figure 2.1 represent (though not exhaustively) that knowledge for the individual in question. On being made to talk about dogs, the person may mention that they have four legs, a tail, that they bark. The person may explain that a dog is an animal, or, if he or she is scientifically minded, a mammal. He or she may also describe the procedure "taking the dog for 'walkies'".

Figure 2.1: A Person's Concept of "Dog"



It is possible (indeed, highly likely) that no two people will share exactly the same set of connections to the notion "dog", and they therefore have different "concepts of dog", or, more correctly, possess the concept to different extents. This need not imply, though, that their concepts of "dog" are incompatible or subjective. On the contrary, their meanings for "dog" are in the public domain and they will normally be able to agree in their application of the concept in day-to-day interaction. Difficulties might arise, though, if the term "dog" were to be used in a different language-game. Such might occur if, for example, the term is used in its more specific sense of "male canine", or in the expression "dogfight". In these situations, new sets of connections have to be assumed for communication to succeed.

2.2.7 Everyday Constructs and Scientific Constructs

In the preceding section, references were made to how, within particular language-games, a scientist might use the word "dog" in subtly different ways from a lay person. It is possible to extend the argument from this simple example to the case of an abstract scientific construct, such as "energy". Here, scientific and everyday uses of the term might differ quite radically. The lay person will most likely be able to use the term in a set of idiomatic expressions to do with "running out of energy" and "being full of energy". These would form part of the person's concept of "energy" and would tend to encourage thinking of energy as a fluid, like a fuel, which can be consumed (see the discussion in 2.2.5). The expert, on the other hand, although probably sharing the lay person's uses, will also be able to use the term in expressions like "energy is conserved" and "changes involve energy transfer". The expert would probably regard the connections evoked by these uses of the term "energy" as central to its scientific meaning. Both of these people would have a concept of energy, but their ability to apply it across a range of contexts would differ substantially. The advantage that the scientist has is not that she or he lacks the various "lay" meanings of energy, but rather that she or he is able to identify contexts in which they are inappropriate.

Scientific constructs derive their meaning from sets of relationships in a public-knowledge theory system. This theory system differs from those used in day-to-day conversation in that it gives greater range of explanation, is consistent with a wider range of empirical evidence, and is probably more logically consistent and less *ad hoc* (see 2.1.1). It is to this theory system that children need to be introduced in their learning of science. But just access may not be enough. They also need to appreciate that this theory system has utility in addition to the existing theory systems which they have used and found true on numerous occasions. This implies, firstly an understanding that the power of scientific theory lies in such attributes, and secondly the capacity to use the terms appropriately in meaningful and purposeful communication.

2.3 A Summary

In the present chapter, it has been shown that the activity we call science consists in exploring the nature of the universe from *within* an evolving theory system. This theory system is a set of relationships between constructs, maintained and developed through activity and discourse (language-games) within the scientific community. Science progresses through developing theories that are able elegantly to subsume more and more explanations of the way the world is, and expert personal knowledge of science involves adopting a similar hierarchical structure to organize what is known. By contrast, children, as novices, typically adopt conceptions that are limited in scope, and rather more contextually bound.

Because of its theory-dependence, science does not provide a uniquely true account of the universe; there are other possible theory systems, each itself arising out of and sustained by a community of discourse. Learning science was characterised in this chapter as involving gaining access to scientific meanings, which itself implies operating within an appropriate community of discourse. Like science itself, learning science can never be an entirely private affair.

The "alternative conceptions" movement has, historically, been guided by a primarily individualistic psychology of learning. Such approaches typically lack an essential perspective, that of the socially determined nature of meaning. It is through supplying this perspective that the link is forged between public and personal knowledge. For while individuals must actively be involved in making sense of the situations they experience, they do so using shared resources: a working language with which they can communicate, however partially; other individuals with whom they can interact and who will be the measure of the success of their communication; and a relatively stable world to which consistent reference can be made in the course of this shared activity.

Although language is used differently in different language-games, the very fact that communication is a feature of meaning-making means that, in principle, any language user can gain access to the theory system being applied by another. Theory systems can "talk to each other". And yet there is a paradox here, for if understanding science implies working *within* its theory system, how can one enter that theory system without understanding it? In the next chapter, the theoretical background developed here will be applied to the practical task of resolving that paradox for the purpose of learning science in school.

3

LEARNING SCIENCE IN THE CLASSROOM

My main thesis ... is that the learner should take more part in the formulation of knowledge. If this is accepted then the small group clearly offers a way of distancing the teacher's control. (Douglas Barnes, 1976, p.191)

3.1 The Nature of the Task

In the previous chapter, a theoretical background was developed, against which prescriptions about the teaching and learning of science in school could be located. It was shown that personal knowledge is not constructed laboriously from first principles by each individual, striving alone to make sense of a continuous stream of (intrinsically meaningless) stimulation. *Personal* construction of meaning depends, not simply on the recollection of personal experience, but also on the underlying, and necessarily prior, *shared* meaning system. Each individual is born into a world that already makes sense, by and large, to significant others, who proceed to draw the neonate progressively into their patterns of meaningful communication. Moreover, it was pointed out that the underlying meaning system does not function as some monolithic construction, but rather as a diverse collection of more or less specific language-games. When participating in any one language-game, some elements of cognitive structure, as it were, jump into focus at the expense of others.

Children in school already operate successfully within a wide range of language-games before they encounter science teaching. However, this very success can lead them into mis-classification of phenomena, when contrasted with the specialized theories of science. The crux of the matter, then, is how to give children access to the language-games of science. It is assumed, in addressing this question, that this implies language-and-practice; not simply a grasp of the words of science, but also of the way of thinking and acting that is intrinsic to using those words.

3.2 Approaches to Teaching: A Review

In this part of the chapter, two teaching approaches will be examined that could be offered in response to this need. The manner in which this is to proceed will be to offer up each approach as though necessary and sufficient in itself (as thesis and antithesis, as it were), whilst recognizing that probably no educationist would accord either such unique status. The purpose in doing so will be, not to incinerate a "straw person", but to contrast the two approaches by throwing them into sharp relief, and to demonstrate how they may in practice complement one another (synthesis).

The two approaches differ in respect of their philosophical perspectives. The first advocates confrontation and rejection of misconceptions, and seems, if taken through to its conclusion, to imply the point of view of metaphysical realism (Putnam, 1981). On this view, it is necessary only to check the correspondence of our conceptions against the world for us to see the error in our ways and exchange our beliefs for those that mirror the world accurately. Thus meaning construction is attributed primarily to the individual in interaction with a physical world that affords only one true description.

The other approach takes the internalist perspective that *all* contact with reality is mediated through our conceptual system (*ibid.*). On this view, more than one true account of the world is possible (though not just any account). Progress is seen in terms of adding to the range of language-games in which we can participate together with developing the ability to choose appropriately between them. Meaning construction is seen as located in successful communication between individuals as they participate more and more in a shared a form of life.

3.2.1 A Confrontational Approach

Much of the literature directed at the problem of children's alternative conceptions features the following broad approach, here taken from Driver, Guesne & Tiberghien (1985):

- the pupils make their own ideas explicit;
- discrepant events are introduced, which contradict the expectations generated by the views pupils have made explicit, leading to conceptual conflict;
- "Socratic" questioning is used to help pupils identify lack of consistency in their views, and therefore to encourage restructuring of their ideas;
- the pupils are encouraged to generate new, meaningful conceptual schemes, which they evaluate;
- the new ideas are applied in a range of situations.

The first difficulty with this approach is that it depends on the direct confrontation and refutation of children's existing concepts. The implication is that, once faced with an event for which they cannot account, children will be able recognize this as a contradiction in their thinking, and accept the need to develop a framework of understanding *at a higher level of generality* to reconcile the conflicting possibilities. But is this likely?

An assumption that seems to lie behind this approach is that by working in this way, children are adopting scientific methods and rationality. Glynn, Yeany & Britton (1991) acknowledge this parallel explicitly. There are two points to make in response. Firstly, children's theories do not necessarily meet the requirements of the version of scientific rationality implied, and secondly, that version is in any case flawed.

Falsification, as advocated by Popper (1970), requires theories to be stated in a form that facilitates refutation, and severe attempts to be made to achieve that refutation. But there is no guarantee that children's theories can be so expressed. Further, this form of falsificationism is too strict (Lakatos, 1970). The point was made in 2.1.1 that a discrepant observation does not overthrow a theory. Instead, anomalies accumulate, and there are attempts to accommodate them by adjusting the theory's "protective belt" (*ibid.*). A theory falls into disuse only when seen in relation to a succession of theories with increasing heuristic potential and, crucially, when the successor gains acceptance in the scientific community (Lakatos, *op cit.*; Kuhn, 1970b). (Even this is not always the coup de grâce, though, as the case of Newtonian mechanics' persisting in the company of Relativity theory demonstrates.) So falling into disuse is a social, rather than an individual, phenomenon. There is no basis for assuming that an individual will come to reject a conception without *first* having encountered and understood a potential successor that offers the possibility of progress. Added to this, there is evidence that children do *not* necessarily feel the need to develop theories of high generality in preference to limited-range theories, including evidence presented in the same volume as the above steps to conceptual change (see also Osborne & Wittrock, 1983).

So in relation to the penultimate of the above steps to conceptual change, there are a number of possible outcomes. Pupils may simply create new, situationally specific, solutions. Or the children will construct new, more general theories that still do not align with those of science. Or they will somehow stumble upon just the right set of constructs to enable them to adopt the scientific ideas. It is not clear how the former two possibilities are in any way helpful, except perhaps as exercises in theorizing. It is also not clear how the latter (though logically possible) is likely, except by the most extraordinary coincidence. For as Kuhn (1970b) has observed, all scientific theories contain an element of arbitrary commitment.

It is noteworthy that Driver *et al.* (1985) do not explicitly suggest offering pupils aspects of scientifically acceptable theory. Their emphasis is very much on generating a range of possible alternatives, and especially on the pupil's rôle in this. Other writers (such as Cosgrove & Osborne, 1985, and more recent work by Driver *et al.*, 1994) make more explicit recommendations about introducing appropriate scientific ideas, but the emphasis on generating alternatives remains. The assumption seems to be that each child should construct *personally* the correct scientific theories *de novo*, a view that is shown in this thesis to be incoherent.

It is worth sounding a note of caution at this point regarding the notion of "Socratic" questioning, as it is in response to this that pupils are expected to generate valid scientific ideas. The basis of Socrates' supposed success lies in the belief that all the ideas we are to acquire are already in place in our minds, though inaccessible, and that "learning is nothing but recollection" (Priest, 1991, citing Plato's *Phaedo*). Hence the skilled user of this technique is setting up situations that "trigger" the appropriate ideas, bringing them into consciousness. Such beliefs are contrary to those underlying this present thesis, turning the Socratic process into an attempt to construct knowledge *ex nihilo*. It was pointed out in the preceding chapter that the origins of our concepts lie primarily in the language-and-practice of the community in which we find ourselves. Scientific theories are the result of considerable and cumulative intellectual effort within one such community, and it is with this that the learner must engage.

In pure form, then, the "confrontational" approach to learning science defies both logic and practicality. Key aspects are shown here to be in difficulty. That there is a place both for challenging existing ideas and for discussing some viable alternatives need not be in doubt. But the purpose of challenging existing ideas should be seen as identifying the limits of their applicability vis-à-vis a viable alternative. For above all there must be the possibility of children's coming to share in an accepted, coherent theory system that will enable them to take their thinking forward, and it is on this crucial point that no guidance is given. Where does this better theory system come from? Constructivist approaches to learning such as that outlined by Driver *et al.* (1985) take as axiomatic that learners must construct their own understanding. But this may have one of two senses. If it is taken to mean that learners should personally invent their own theories through inductive reasoning based on empirical data, then it is flawed. Children cannot be expected to discover just those theories that the scientific community has adopted as the most fruitful. If, on the other hand, it is taken to mean that learners should go through a conscious and deliberate process of linking ideas together in their own thinking, and of checking their understanding for coherence and consistency as well as against empirical evidence, then it is entirely in keeping with the basic premises of this present thesis. For this leaves open the possibility of introducing scientific constructs directly into the language-and-practice of the classroom. In later publications, Driver *et al.* (1994) have endorsed the latter view more explicitly. Next we shall explore what such an approach might be like.

3.2.2 An Apprenticeship Approach

Driver *et al.* (1985) suggest, it appears quite correctly, that learning to work with scientific theories should be seen as a long term aim. The move towards expert knowledge in science will consist in the gradual expansion and reorganization of the personal network of interrelated concepts. But as yet we are no closer to knowing how to effect an entry into this highly specialized network. This is the problem of the "hermeneutic circle", in which we must approach a strange "text" (the scientific meaning system) from the point of view of our existing understanding, and engage it in something like a conversation (Rorty, 1980; Sainsbury, 1992; Eger, 1992).

A conversation involves an interplay between at least two viewpoints. Underlying this discussion is the assertion that personal understanding depends on concepts and connections already present in the learner's cognitive structure. One of the interlocutors in this "conversation" must therefore be the set of concepts and connections that the learner already possesses. The other partner in the exchange is the network of connections that comprises the relevant public-knowledge theory system of which we wish to improve the pupil's grasp. It is not that one partner comes, as it were, to dominate the other, for that would not be a true conversation. Rather, a common ground is established between the two viewpoints: "an understanding of an individual theory is a matter not only of understanding the links in the new theory but also of grasping its place in relation to the other theories of the total system" (Sainsbury, *op cit.*, p.120). However, children may need help to make the appropriate links between these new ideas and their existing conceptions. Here, care is needed from the outset to ensure that the new concepts are developed in a way that is consistent with scientific meanings. At this stage, the foundations may be laid for children to come to appreciate the power of scientific theory as a meaning system. This meaning system must be applied both in discourse *and* related practical activity, so as to establish a specifically scientific language-game. In doing so, the logical coherence and wide applicability for these new meanings should become clearer.

Throughout, it is important to recognize the necessity that meanings be shared and communicable. Novak & Gowin (1984), whilst seeking to promote a view of learning similar in many ways to that described here, have overemphasized the rôle of the individual in achieving that learning: "learning is a responsibility that cannot be shared" (p.6). Whilst it can hardly be denied that deliberate attention and effort are required on the part of pupils, such a statement surely undervalues the rôle of the classroom community, both teacher and children, in establishing the shared meaning system. If the essence of meaning is that it must be communicable, then this can *never* be an entirely individual responsibility.

The notion of a community that works together to share the task of constructing understanding, yet in which some members are clearly more knowledgeable than others, characterizes what may be called "apprenticeship" approaches to learning. These receive warrant from two complementary sources. The first of these is the meeting of perspectives advocated by Sainsbury, who says, "by a 'meeting of perspectives', I mean the process by which each participant in the conversation comes to understand the relevant theories, the set of conceptual connections, being applied by the other" (*op cit.*, p.113). This "represents a greater degree of participation in the theory-system, in the form of life" (*ibid.*, p.114). The second comes from a school of psychology originating in the Soviet Union. In the next section, it will be shown how this can throw new light on the question of how children can make progress into the scientific meaning system.

3.2.3 The Socio-Cultural Perspective

The ideas presented here derive from Soviet psychologists amongst whom Vygotsky and Leont'ev were influential, and have subsequently been taken up in the west by Bruner; by Wertsch; by Newman, Griffin and Cole and their collaborators at the Laboratory of Comparative Human Cognition; and by Vedder, amongst others. Central to this "socio-cultural" viewpoint is the idea that ways of knowing are situated in contexts and culturally constructed (O'Loughlin, 1992).

In an educational setting, content and goals are culturally determined, and in general the pupil's learning experiences are regulated by other persons, that is, teachers (Vedder, 1985). This cultural dimension is what is distinctive about human learning. It is in interactions between persons that we first learn the procedures that constitute thinking, in the human sense. Ultimately, according to Vygotsky (1978), these procedures become internalized (by some little understood process), a move characterized as from other-regulation to self-regulation. An essential element in this view is that of a "tool" available in the culture, of which linguistic signs are important examples. According to Leont'ev, the function of a tool is not discovered through unaided exploration but through involvement in activities in which it is used (Newman, Griffin & Cole, 1989). In almost every respect, these ideas parallel the Wittgensteinian view developed so far in this thesis, as Edwards' & Mercer's (1987) summary shows:

> Vygotsky was proposing that children's understanding is shaped not only through adaptive encounters with the physical world but through interactions between people in relation to that world - a world not merely physical and apprehended by the senses, but cultural, meaningful and significant, and made so principally by language (p.20)

A key construct within the socio-cultural framework is that of ***socio-cognitive conflict***. This is "a conflict between persons who handle the same task in a different way because they experience the task differently, but are still striving for the same solution" (Vedder, *op cit.*, p.36). It is best conceptualized as a perceived communication breakdown, occurring when participants construe the same action differently. It need not be major, and need not involve argumentation. The desire to maintain communication motivates a "meeting of perspectives". An extreme example of this would be that of an expert (such as a teacher) cooperating in a task with a novice (a pupil). The teacher's responsibility is to explain the expert view to the child, so that the child can share the teacher's interpretation of the task and complete it successfully. The pupil's responsibility is to explain his or her perspective to the teacher; the conversation must have two sides. Pupils working together on a task may also have differing perspectives, and so need to engage in a similar process of "negotiation of meaning" in order for communication to succeed. This negotiation takes place in interactions between the parties involved, and is mediated chiefly through language.

The negotiation process can be thought of as operating on two levels. When faced with the task of explaining to others what he or she thinks about a topic, a learner needs to express his or her existing understanding clearly. This would involve clarifying the meaning that the topic already has for the learner (that is, specifying the connections that are being applied). It may entail adapting the explanation to take account of the recipient's level of understanding and experience (King, 1990). This results, firstly, in the "pooling" of available knowledge on the topic, and, secondly, in revealing to participants the level of their own understanding of the topic. These in turn may lead to restructuring personal knowledge in order to communicate effectively with others, or to resolve inconsistencies that arise in the course of this communication.

Reflective awareness and evaluation of ones own understanding have been subsumed within the construct "metacognition" by some writers (such as Garner, 1990), although this term is not used consistently in the literature (Brown, 1987; Prawat, 1989). Prawat contrasts "tacit" or "unanalyzed" knowledge with "explicit" or "analyzed" knowledge. The former is understood only superficially, and can only be applied in a routine manner. The structure of such knowledge is not known to the learner, and not therefore available to be operated upon. However, if a strategy were available that could enable the learner to access and analyse tacit knowledge, this might assist the learner to integrate such knowledge more fully into cognitive structure. Prawat (*op cit.*) claims considerable support for the view that verbalization (and in particular, communication to others) is the most effective means of making tacit knowledge available for examination. Strategies proposed to facilitate socio-cognitive conflict have therefore typically emphasized structured discussion among pupils (Light & Glachan, 1985; King, 1990).

3.2.4 The Growth of Ideas in Social Settings

As with the confrontational approach described above, there is a parallel between the way scientific knowledge is constructed and the way individual learning proceeds in an apprenticeship. However, the parallelism is rather different.

Sociologists of science, such as Knorr-Cetina (1981), have drawn attention to the way that both the social and the physical environments are influential in shaping the creation of scientific knowledge. In her analysis, the ideas put forward by scientists are characterised by a degree of inexactness or "indeterminacy", which, rather than being disruptive of progress, actually facilitates it. This is "a necessary prerequisite for progressive, organised adaptation, and thus for survival and reconstructive change" (p.10), much as mutation is in biological evolution. On this view, ideas may be introduced by individuals, but it is the scientific community that serves to "select" or modify those ideas, making the growth of scientific knowledge very much the result of "discursive interaction".

Newman, Griffin & Cole (*op cit.*) have described very similar processes at work constructing knowledge in classrooms. The fact that an utterance can take a rôle in more than one frame of reference (language-game) means that participants in a conversation have "room for manoeuvre", thereby facilitating change. In the so-called "zone of proximal development" (a construct introduced by Vygotsky), understandings are achieved, through interaction, that individual participants could not achieve unaided. This happens when a participant "appropriates" the idea of another, though not necessarily in the sense intended by the other, and incorporates it into the continuing discourse. Thus the discourse can take on a complexity that is not due to any individual's contribution alone.

3.2.5 The Apprenticeship Approach Examined

In contrast with the confrontational approach outlined earlier, this approach does not assume that an existing conceptual structure must be shown to fail before a scientifically acceptable set of meanings can be learned. Instead, the assumption is made that normally the shortcomings of a theory *only* become fully evident with reference to a more powerful rival. Confrontation, if it occurs, is a *result* of, rather than a precursor to, fuller participation in the scientific meaning system.

That alternative conceptions may be abandoned by pupils in the light of "better" theories without their being explicitly challenged has received tentative empirical support. Samarapungavan (1991) carried out a study with children in the age range six to eleven years to determine their ability to apply principles of scientific rationality in choosing between competing theories. The four principles chosen for study (range; non-ad hocness; empirical consistency; logical consistency) are very similar to those proposed by Putnam (1981). The findings suggest that the children were able to apply criteria based on these principles to select between the two alternative theories they were offered. Two further aspects of the findings are of particular significance here. Children (especially the younger ones) were more successful in applying the criteria when they found the "better" theory plausible. More interesting, though, is the finding that children

> showed a systematic preference for theories that could account for a broader range of observations even though the narrower theories were not directly disconfirmed by the observations they could not handle (p.44)

and that many of them could justify their choice in these terms. Although this was a "clinical" study, with the initial choice of competing theories made by the researcher, the findings suggest that understanding the relationships between a set of constructs and appreciating their power in "explaining" reality may be factors in children's adopting them, even if a competing set of constructs has not been "refuted" directly. Muthukrishna *et al.* (1993), working in more naturalistic classroom contexts, but with pupils aged around 14, reported that "conceptually integrated" instruction in a science topic was successful in displacing alternative conceptions without the need to address the latter directly. These studies were not set up to test an apprenticeship approach to learning *per se*. However, the findings are entirely compatible with such an approach, and provide it with empirical backing.

Critics of the approach would no doubt cite as counter-evidence the body of research, already alluded to, that documents children's failure to grasp scientific constructs. The problem is, essentially, the "learning paradox": how can a learner acquire a new cognitive structure without first having a more advanced or complex structure to which it may be related (Shuell, 1986)? However, complexity, in the apprenticeship approach, is a feature of interactions external to any one individual. Because applying a meaning system is a public, rather than private, activity, children need not be trapped behind conceptual bars of their own making. Their entry into the new language-game may be slow, partial and faltering, but because it is supported by the community, it can proceed. The paradox is resolved.

We have, therefore, a view of learning science in which children are guided into a new way of talking about the world that changes how they perceive the world. This implies certain conditions, if it is to be successful. The first is that children should be talking *about* something in the world; the developing scientific language must serve to integrate the range of children's scientific experience in school, including practical experience. Practical tasks (valuable in themselves for learning how to *do* science) should also be seen as a context for applying scientific meanings. This relates to the second condition, which is that the teacher should become, not a mere provider of practical experience but one who moves the children on in their interpretation of that experience, through interacting with them. But by the very nature of their organisation, classrooms are quite unlike the settings in which apprenticeship (in its original sense) operates; there is not always the opportunity for teachers to engage children in conversation in quite the depth implied. One proposed solution to raising the potential for interaction in the classroom has been cooperative learning initiatives, in which children engage each other in talk about the subject matter they are studying. In the next part of the chapter, relevant research on cooperative learning is discussed.

3.3 Cooperative Learning Research

A variety of research perspectives have been taken on cooperative learning. A number of studies have tested the hypothesis that verbal interaction between pupils is effective in promoting learning. Amongst these, several were focused on the rôle of specific types of interaction. Noreen Webb has been prominent in this research programme.

Webb (1982a) describes a typical investigation of collaborative learning in small groups. The subjects were grade 7 and 8 students in US high school, and were learning a unit on exponents and scientific notation. Because ability composition of the groups was of interest, the students were assigned on a stratified random basis (with quotas selected from each ability level). They were instructed to work collaboratively on exercises in the unit, and achievement was measured using a teacher-made test comprised of problems equivalent to those in the exercises. Ability was determined with reference to a test of mathematical reasoning developed within the school. Partial correlations, correcting for ability, were calculated between individuals' posttest achievement scores and categories of verbal interaction engaged in during group work. Interaction categories of interest were ***giving explanations*** and ***receiving explanations*** in response to questions or errors, ***giving short-answer feedback*** and ***receiving short-answer feedback*** in response to questions or errors, and ***receiving no answer*** in response to a question or error. "Short-answer feedback" refers to unelaborated responses, such as simply supplying the answer to a problem.

Giving explanations about what was being studied was modestly correlated with achievement[1]. This correlation was slightly higher than that between achievement and receiving explanations, and, though small, suggests there may be an effect due to verbalization. The correlation between achievement and giving short-answer feedback was not significant. On the other hand, substantial negative correlations were recorded between achievement and receiving no explanation.

---

[1] $r = 0.22$, with ability partialled out

On this basis, Webb (1982a) concludes that giving explanations is beneficial for personal learning, and (1982b) reviews other research supporting this thesis. An alternative conclusion, however, would be that it is only those who understand an idea who are able to provide explanations. The fact that ability was removed as a factor tends to favour the former conclusion. However, a weakness of the study is that no pretest information on achievement was collected, and consequently there was no control over the effects of prior domain-specific knowledge.

In a subsequent meta-analytical article, Webb (1989) makes a detailed examination of correlations reported between the above interaction variables and achievement in nineteen published research papers. These covered a wide age range, from US grade 2 to grade 11. Overall, the figures tend to support the earlier findings. Of five studies in which significant correlations were found between receiving explanations and achievement, all were positive, and some were substantial[2]. The effects of other levels of interaction were more varied; however, providing only the correct answer without explanation was almost always (with one exception) negatively related to achievement. Giving explanations was consistently related to achievement. Twelve studies reported on this relationship, and with one exception (not statistically significant), demonstrated a positive relationship[3]. These data add further support to the view that the pooling of knowledge, and in particular the giving of full explanations, within a group is generally beneficial to learning. Applying knowledge in acts of communication seems to improve personal understanding. The age of the participants did not appear to be a factor in whether they could benefit from cooperative learning.

An interesting negative result comes from Vedder (1985). Vedder, in a well-controlled experiment, found no significant difference between the performance of individuals and cooperative groups (aged 8-11 years) on mathematical problems, despite training the latter in cooperative problem-solving procedures. However, a subsequent examination of group interaction indicated that the pupils were not cooperating in the ways intended. This lends further support to the view that the occurrence of

---

[2] $r = 0.21$ to 0.63

[3] Significant partial correlations ranged from $r = 0.22$ to 0.52

particular group processes might be a critical factor in the learning gains reported in most research in this field.

Another focus for research on cooperation has been on the effect of structuring collaborative tasks. King (1990) provided specific training in strategies for cooperative learning from lectures (rather than the more typical problem solving). College level student teachers were involved, who were studying methods of evaluating outcomes from teaching. They were trained in reciprocal questioning, using question stems such as "Explain why ....", "How is .... related to ....?" and "How does .... effect ....?". These were applied in a post-lecture review discussion, which was recorded and analysed using a version of Webb's categories. Understanding and applying the lecture content (achievement) was measured using a short written test. The students trained in reciprocal questioning generated significantly more high-level interaction than control group students who studied the same material, but who were only directed to discuss it. They also out-performed significantly the control group in the posttest. For this, control was exercised (by means of analysis of covariance) over ability of the students to comprehend material presented in lecture format, as indicated by a separate written test. As with Webb's studies, there was no pretest of domain-specific knowledge. However, in this case, it was not the correlation between the interaction variables and individual achievement that was of interest, but the difference in the means of these variables between groups. Although intact groups were used in the study, they were found to be equivalent in terms of ability, and it is unlikely that one group would have had a clear advantage in terms of prior knowledge. A follow-up study, in which structured reciprocal questioning using the above question stems was compared with unstructured reciprocal questioning, produced similar results with (it appears) different student groups (*ibid.*). This adds weight to the supposition that the quality of verbal interaction affects achievement, and not vice versa.

King (1992) has adapted these ideas to individual learning by college students. A similar set of question stems was used (with the notable addition of "What is the main idea of ....?). Another experimental group was instructed in summarizing information, and was shown how to link together important ideas in a topic using their own words, and how to identify the main idea and sub-topics. A control group used more traditional note-taking strategies. Groups were pre- and posttested using a multiple-choice test of comprehension. Both experimental groups performed significantly better on the posttest (adjusted for pretest scores) than the control group. On a delayed posttest, both experimental groups still out-performed the control group, but this difference was only statistically significant for the self-questioning group.

The work of Barnes (1976) takes a different perspective. He suggests, on the basis of qualitative research, that negotiation of meaning is particularly encouraged when groups are given freedom to engage in looser "exploratory talk" before having to produce a more explicit and precise "final draft" outcome for a wider audience. Here, an unthreatening atmosphere is important, distanced from the direct control of the teacher and in which each group member's contribution is valued.

There is, then, a body of evidence that strategies that engage learners in a conversation about what they are learning (either with themselves, or with other learners) are effective in promoting understanding. Processes that seem to be at work are the elaboration of learned material through generating explanations about it, and structuring it by identifying hierarchies of ideas and relationships. To some extent, the effects have been observed with learners of various ages.

3.4 A View of Science Learning

It is now appropriate to draw together the strands that have been explored in this and previous chapters, in order to reconsider what it means to learn science effectively in the primary classroom.

As described in Chapter 1, children are *required* under the National Curriculum to learn about certain of the constructs in science, and it is a basic assumption of the present research that they should learn these with *understanding*. This implies that children should know when it is *appropriate* to apply scientific conceptions, rather than the more restricted but perfectly valid everyday conceptions that they also hold. This cannot happen if children see these different perspectives as interchangeable. This in turn implies therefore an emergent appreciation of the rationality for scientific progress, that is, that for certain purposes scientific ideas represent progress vis-à-vis everyday conceptions because of their comprehensiveness, functional simplicity, instrumental efficacy and coherence. These are long term aims. The task here is to identify ways in which progress can be made towards them during children's time in the primary school.

The growth of personal knowledge was characterized above as resulting from a "conversation" between a pupil's prior concepts and the constructs of science. It consists in the development of more and more conceptual connections, giving increasing scope and precision to personal meaning. The result of this growth of the personal perspective should be evident in children's ability to apply scientifically acceptable meanings in appropriate activities. The direct confrontation of "alternative conceptions" may or may not play a part in this. What does not make sense is the idea that children can construct a more successful theory on the sole basis of reflecting on the empirical refutation of an existing one.

Once children have begun to identify connections between constructs relevant to a scientific domain and to see how these can be used to refer to the world, they have broken into the "hermeneutic circle" and made a start in assembling a coherent theory structure. Having made this start, they should then be helped to relate this structure to each new experience of the domain, thereby increasing its range of application and elaborating further the network of connections. As a consequence of this process, the scientific ideas may become more plausible, and increasingly well-corroborated.

The approach to achieving "conceptually integrated" learning that is implied here is "constructivist" in the sense that children should personally be active in making the necessary conceptual links. But this construction must be constrained by the need to embrace culturally accepted scientific meanings. What children construct must withstand the test of public scrutiny, and construction cannot be an entirely private affair. It must involve an attempt to coordinate personal understanding on the part of each pupil with the meanings being applied by the others and the teacher: there is to be a "meeting of perspectives".

To achieve this, learning activities are needed that reflect and promote coherence and range of explanation in children's developing cognitive organization. In such activities, meaning, as connections and relationships, would develop through the successful application of concepts in communication. Such an approach is necessarily interactive: whilst individual contributions will certainly be a feature, so too will appropriation and "natural" selection by the group of these contributions.

It is now possible to suggest how concept mapping can play a part in this meeting of perspectives. Some procedure is required that will prompt and assist pupils to coordinate knowledge and experience within a domain, and focus discussion about the meaning of new information and how it relates to what they already know. This discussion should engage pupils in giving and receiving explanations using the constructs being learned and the links between them, rather than in rehearsing information verbatim. By externalizing the processes of analysing knowledge, linking in new knowledge and restructuring knowledge to accommodate new information, the learner should be encouraged to be explicit and accurate, so that what is communicated complies with public criteria. Concept mapping appears to be a technique that meets these requirements, and will be described in the next chapter.

4

LITERATURE REVIEW

The making and remaking of concept maps and sharing them with others can be seen as a team effort in the sport of thinking. (Joseph Novak & Bob Gowin, 1984, p.19)

4.1 Introducing the Concept Map

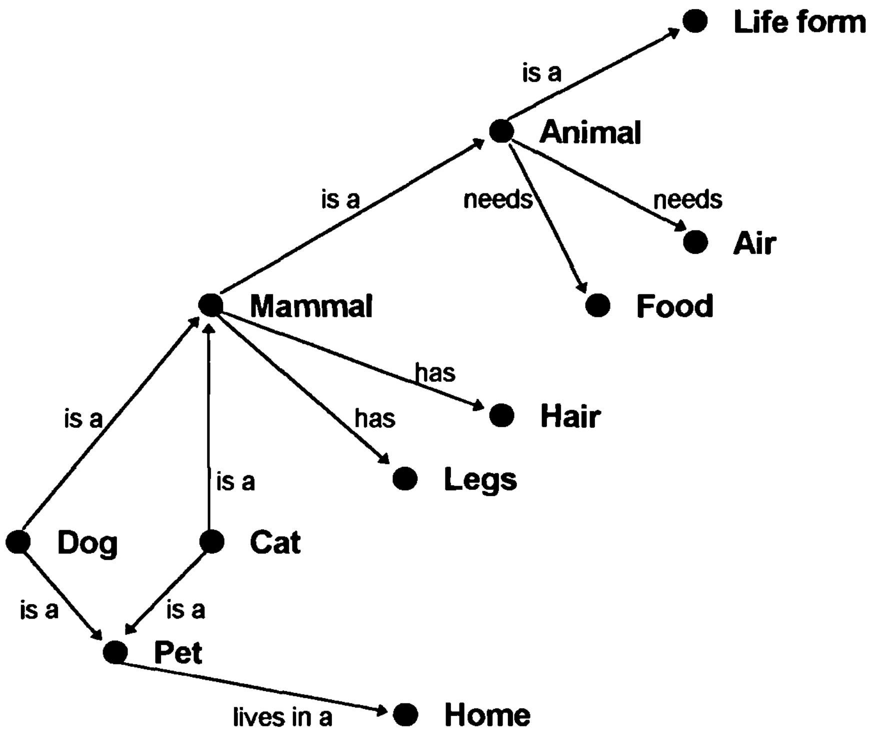
Concept Mapping has been proposed as a "metacognitive tool" with the potential to improve meaningful learning by Novak (1990a). The basis for this claim, and relevant research, will now be reviewed.

4.1.1 The Origins of Concept Mapping

The procedure known as concept mapping emerged from Cornell University in the early 1970s. It seems clear, however, that it owes much to similar ideas in use before this time. The result seems to be a synthesis of these ideas, rather than a novel conception.

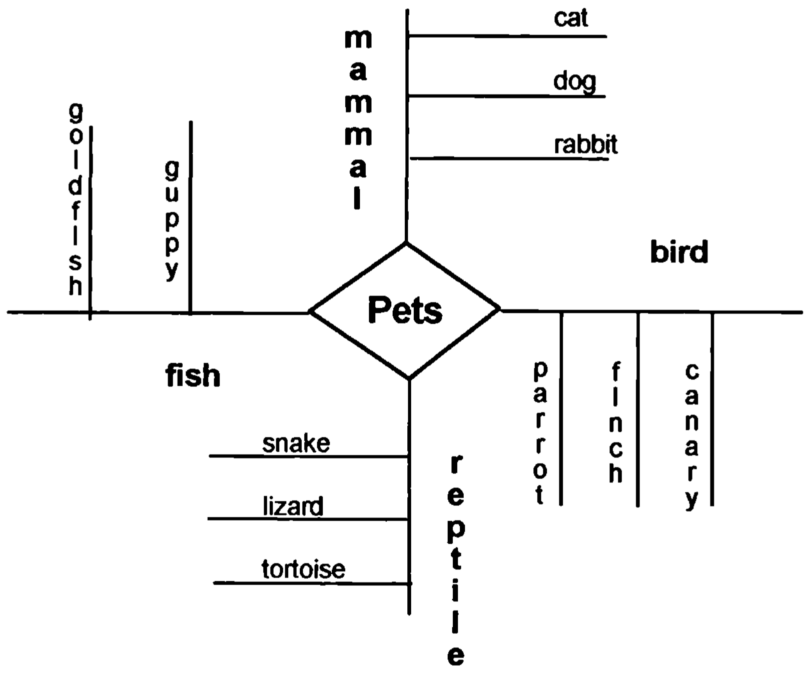
A diagram identical in form to a concept map was used by Collins & Quillian (1969) to illustrate the idea of a "semantic network". Its purpose was to illustrate the manner in which concepts are related in semantic memory, and its features reflect those of the "Spreading Activation" theory, which formed the basis of the paper. Concepts are represented in the diagram as "nodes", or points, joined to related concepts via a network of "pointers". Each node is identified by a dot, and is labelled with the name of the concept. Each pointer is labelled with a specification of the relationship between the concepts it connects. The pointer is directional, and can only be "read" one way. The resultant diagram is therefore able to represent knowledge in the form of a network of propositions. Figure 4.1 shows an example.

Figure 4.1: A Semantic Network ***(After Collins & Quillian, 1969)***



Hanf (1971) describes a means of diagramming the structure of meaning in passages of text. This type of diagram is claimed to be based on Gestalt principles, although the basis for this is not elaborated. The focus consists of an enlarged node representing the concept of central concern, and further related ideas branch from this. Its purpose is to improve understanding and retention of the main ideas in a passage, as part of a reading and reviewing sequence. Figure 4.2 shows an example. Buzan (1974) presents a very similar style of diagram, which he terms a "brain pattern". Among the uses suggested by Buzan are planning for talks or papers and taking lecture notes.

Figure 4.2: A Map of the Idea "Pets" ***(After Hanf, 1971)***



Novak (1990b) locates the origin of concept mapping in a twelve year longitudinal study of children's understanding of science concepts. During the course of this study, large quantities of interview data were collected. A need arose for an economical means of representing these data which would enable the researchers to compare their subjects' cognitive structures before and after teaching. Rowell is credited with using a concept map, in 1973, as a "template" for analysing interviews (Novak & Gowin, 1984; Stewart, Van Kirk & Rowell, 1979). First, a model map representing the knowledge domain of interest (as viewed by the researchers) was constructed by Rowell, and then each interview transcript was examined for the presence of the concepts and propositions in the map.

4.1.2 The Relationship between Ausubelian Theory and Concept Maps

The Cornell research was guided by Ausubel's assimilation theory of cognitive learning (Novak, 1977; Ausubel, Novak & Hanesian, 1978), and the structure of concept maps has evolved to reflect this (Novak, Gowin & Johansen, 1983). Ausubel's theory is highly detailed, and will only be discussed to the extent that it can contribute to the theoretical position established above.

The theory is centred on personal knowledge, which is represented in the theory as a hierarchical network of concepts and propositions. For learning to be meaningful, new concepts and propositions must be linked into an existing structure. To reflect this theoretical position, concept maps are now constructed so as to emphasize the hierarchical structure of the domain. Links between concepts in the map are also now labelled to make the content of propositions explicit. These features were lacking in earlier concept maps (Novak, Gowin & Johansen, *op cit.*), and also the diagrams described by Hanf and Buzan. However, Collins' & Quillian's original semantic network represents an intermediate position. Their network was hierarchically structured, and links were labelled, but only a restricted range of relationship types was included. Since Quillian's semantic network pre-dates the less sophisticated early concept maps, it seems that the latter must have a separate genesis. Also, because the early concept maps lacked Ausubel's emphasis on hierarchical structure, they probably were not derived directly from Ausubelian theory either. Rather, they seem to have started off as a pragmatic solution to a particular problem, with the features of semantic networks and Ausubelian theory grafted on at a later date. The theoretical basis for concept maps, as presently conceived and described by Novak & Gowin (1984), will now be examined.

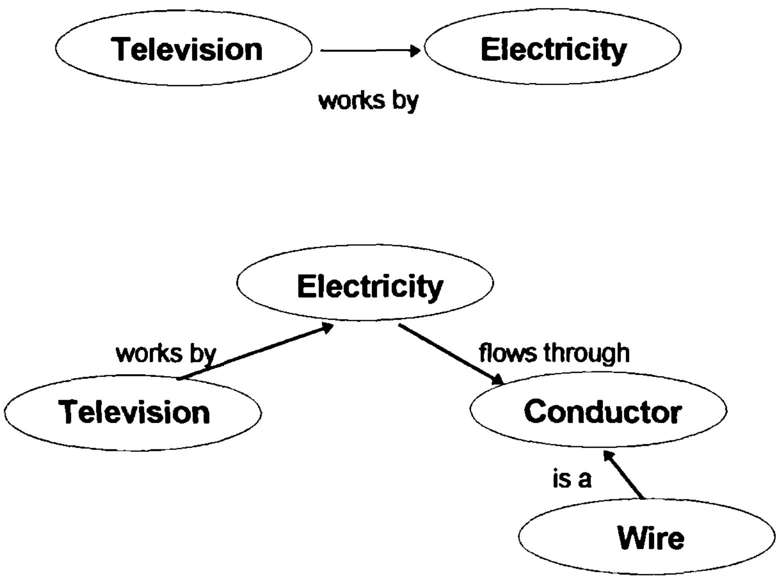
4.1.3 A Psychological Basis for Concept Mapping

For Novak, (1977), concepts are "inventions of man used to describe observed regularities in events", and are designated by a label (p.454). By the term "label" is meant a culturally recognized verbal symbol: a word or words. Novak thus places personal concepts firmly in a public knowledge context. In Ausubel's view, most learning in educational settings is concerned with acquiring the meaning of culturally accepted, verbally defined concepts. Seldom are original concepts generated independently by learners (Ausubel, Novak & Hanesian, 1978).

For Ausubel, learning is meaningful to the extent that new knowledge is related consciously by the learner to relevant elements already in memory (*ibid.*). In meaningful verbal learning, new propositional links are made to existing concepts, so not only is there an increase in the amount of knowledge stored, the meaning of each existing concept is also changed. Concept maps are diagrams that are intended to represent the meaningful relationships between concepts in memory. There are parallels between this position and the Wittgensteinian view presented in the preceding chapters, but the main emphasis here is on the links in the mind of the individual, rather than in the public theory system. Hence most research into concept mapping has focused on individual outcomes.

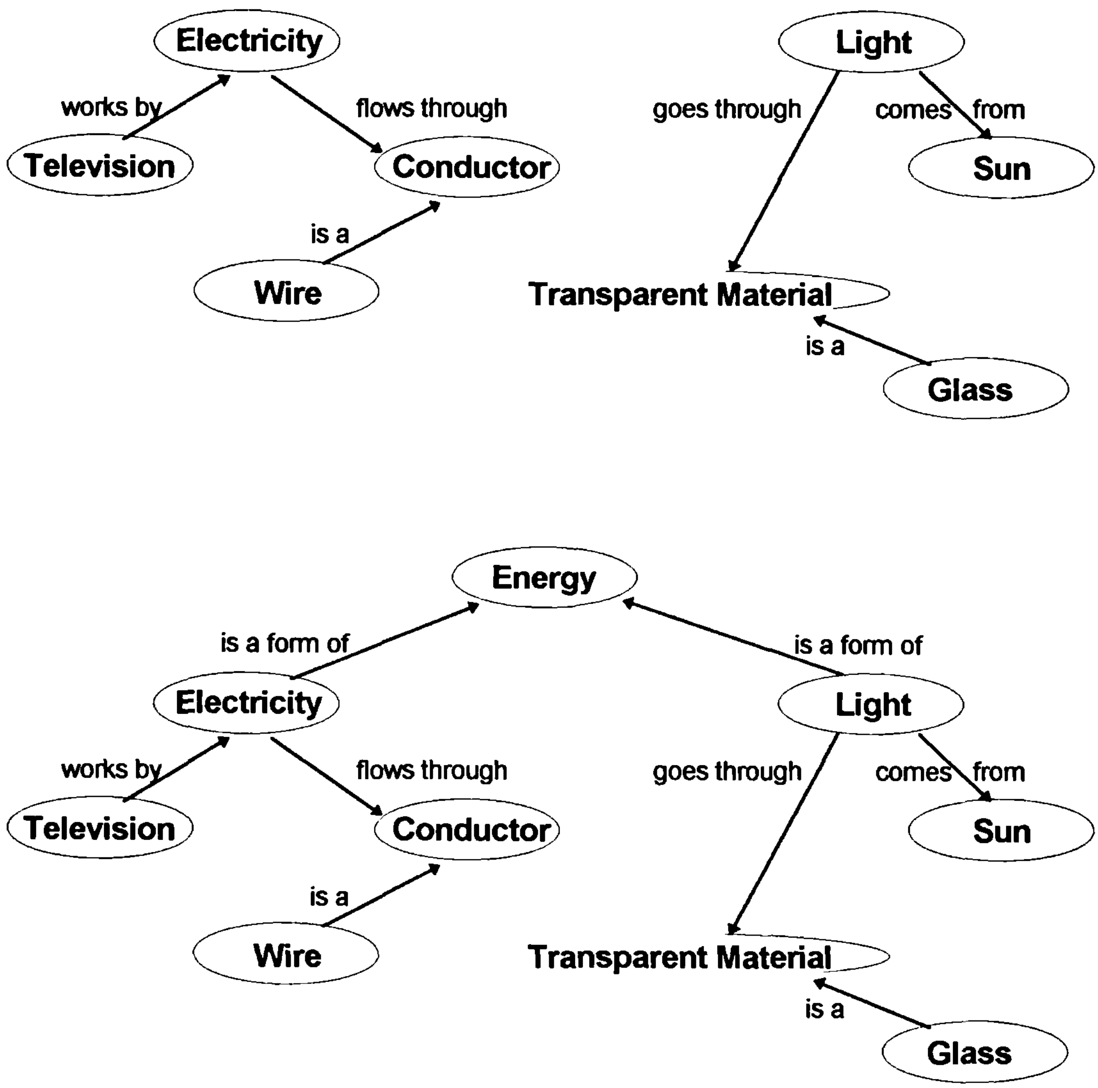
In a concept map, a concept is shown as a labelled node, and a proposition is represented as two or more of these nodes connected by a line (as in the top part of Figure 4.3). In contemporary versions of concept maps, these lines are labelled, and often drawn as vectors to indicate the direction in which propositions should be "read".

Figure 4.3: Concept Maps and Subsumptive Learning



Ausubel distinguishes between three kinds of propositional learning, of which two are represented in concept maps. The first of these is ***subsumptive learning***. This consists of relating new propositions to an established "anchoring" concept (or proposition). The subsuming concept is conceived of as occupying a superordinate position in relation to the new ideas, giving knowledge a hierarchical structure. As subsumptive learning continues, the subsumer is said to undergo ***progressive differentiation***. This denotes a progressive refinement of meaning, through the assimilation of new information to an existing configuration. In Figure 4.3, the concept of "electricity" has undergone differentiation by the subsumption of "conductor". In order that they may represent this feature of knowledge clearly, concept maps also need to have a hierarchical structure. By convention, this is assumed to mean having a top to bottom format, though there is no compelling reason why a central node might not function just as effectively as the subsumer.

Figure 4.4: Concept Maps and Superordinate Learning



Superordinate learning is an important but rare and difficult form of learning (Novak, 1983). It occurs when an entirely new concept is added to cognitive structure in such a way as to subsume existing concepts, through the new propositions formed. As a result of this, existing ideas that were previously seen as unrelated, or even in conflict, may become connected in a meaningful way. This is termed ***integrative reconciliation.*** In Figure 4.4, the new superordinate concept "energy" has been incorporated into memory, resulting in integrative reconciliation between the concepts "electricity" and "light", and also in differentiation of both concepts. Note that the structure of knowledge over the series of examples becomes progressively hierarchical as the more powerful concepts are emphasized. Integrative reconciliation may also result from learning that two concept labels denote the same regularity.

As the above examples show, concept maps are capable of representing each of these two types of learning by way of changes in their propositional structure, progressing towards the kind of conceptual integration characteristic of expertise. There seems to be a *prima facie* case for accepting concept maps as valid representations of knowledge structure. However, there are respects in which concept maps run into difficulty. Novak treats the linking words between concept nodes as being different in kind from the concepts they connect to. Yet the meaning of a word resides in the connections that are be made to it. This implies that linking words, too, are labels for constructs which derive their meaning from a network of relationships to still further constructs. A limitation of concept maps, then, is that they can only give a relatively "coarse-grained", and thus selective, picture of the structure of knowledge. This need not be a disadvantage, though, as a more complex model is not necessarily a better model (c.f. Johnson-Laird, 1985). An important rôle for concept maps in learning is to impose structure and to help learners take a more holistic view of the topic.

Ausubel's third type of learning, ***combinatorial learning,*** is not discussed by Novak in relation to concept mapping. Unfortunately, it is particularly difficult to untangle exactly what Ausubel means by the term, or to consider its relevance to the present discussion, as he declines to give examples.

4.2 The Role of Concept Mapping in Meaningful Learning

Even if it is accepted that a concept map is a valid (though simplified) representation of knowledge, what basis is there for supposing that this can be used to assist learning in science, as it has been discussed above? Central to this section is the notion that the construction of a concept map is an *act of communication*. Thus the *process* of concept mapping may be viewed from the perspective developed in Chapter 3.

It may be hypothesized that the decisions a learner makes when constructing a hierarchical concept map are central to the process of learning. Important decisions would include: "Which is the most important idea here?"; "How does .... relate to ....?"; "What is the difference between .... and ....?". These questions are isomorphic with the generic questions used by King (1990; 1992) to guide learners in reciprocal- and self-questioning strategies (see 3.3). Being intrinsic to the task of constructing a concept map, these strategies can be introduced in the context of a meaningful activity, rather than as abstract skills.

4.2.1 Negotiation of Meaning

To agree on the structure of the map is to agree on the meaning of the constructs represented. When two individuals who understand a domain in different ways collaborate in constructing a concept map, socio-cognitive conflict is likely to arise between them before achieving a consensus regarding how the final map should be structured and on the meaning of the links to be made. Agreement constitutes a problem to be solved, and the public nature of the product makes it clear what is under discussion and when the final goal has been achieved.

The finished map, or interim stages towards a map, may also serve to focus negotiation between the teacher and the pupil or pupils involved in its construction. The teacher will represent yet another source of socio-cognitive conflict. The concept map, by communicating to the teacher the current state of pupil's thinking, can highlight where thought-provoking questions may be effective in leading pupils' thinking forward. The map can then be thought of as helping the teacher identify what Vygotsky (1978) called the "zone of proximal development".

4.2.2 Metacognition

Socio-cognitive conflict and negotiation of meaning arise when knowledge is brought into the "public domain", as it were. There may also be more direct benefits to the learner through the process of externalizing existing knowledge. Some of this existing knowledge may only be held tacitly. To make this knowledge communicable, it must be expressed as propositions. With these available for inspection, the knowledge structure becomes more readily analysed, and this may facilitate the detection of inconsistencies. For the first time, the learner may become aware of some inconsistency in his or her understanding that previously was masked by the inability to hold all the relevant ideas in mind. McAleese (1985) terms this kind of awareness ***anomalous state metacognition***, and it would appear to be a necessary condition for integrative reconciliation.

As well as helping the learner to analyse the relationships between existing concepts, concept mapping can also help him or her to consider how best to integrate new information meaningfully with what is already known. Some of this new information may be in conflict with what the learner believes. Again, by requiring that the learner make the nature of the relationships between concepts explicit, the act of creating a concept map can encourage the learner to identify, and thus eliminate, inconsistencies.

4.2.3 Superordinate Learning

The difficulty in achieving superordinate learning has already been noted. Nevertheless, acquiring an appropriately hierarchical knowledge structure was identified above as an important feature of expert knowledge, and may also be a contributory factor in appreciating the power and coherence of a scientific way of talking about the world.

Creating a hierarchically organized concept map entails identifying superordinate relationships. As part of the procedure for making such a concept map, learners could therefore be encouraged to identify and incorporate in their maps ideas of high generality in such a way as to subsume a range of relevant existing concepts. Concept mapping should, therefore, contribute to the development of a coherent, highly interrelated and hierarchically organized cognitive structure.

4.2.4 The Social Context of Learning

In the scientific community, knowledge is open to negotiation, and can be rejected or changed, creating more coherent structures of constructs, with greater explanatory power, consistent with the evidence. Science has public criteria by which theories may be evaluated. It also has groups of individuals with personal stakes in particular knowledge claims. These features and processes are common to the construction of a concept map by a group of learners.

In the creation of a concept map, knowledge is made available for public scrutiny. The map is constructed according to certain conventions, and above all, should make sense to readers. These features provide the group with public criteria for evaluating the relationships between ideas. As a means of externalizing knowledge for all to see and discuss, concept maps therefore have the potential to focus negotiation of meaning. If, in the process of discussion, learners are called upon to explain, defend or justify their views, this creates a personal interest in, and commitment to, the outcome. This may demand a reassessment of the evidence for the claim, drawing, for example, on the results of practical experiment or on secondary sources. In these respects, concept mapping not only reflects elements of the National Curriculum programme of study, it may also emulate aspects of the exchange of views and of the construction and justification of knowledge within a scientific community.

Overall, these features appear to give concept mapping the potential to support an apprenticeship approach to learning science, as described in Chapter 3.

4.3 Concept Mapping: A Review of Research

Concept mapping has, though, a wide range of applications. Novak lists the following uses (Novak & Gowin, 1984, p. 40-49):

- exploring what learners already know;
- "roadmapping" a learning route;
- extracting meaning from textbooks;
- extracting meanings from laboratory, studio, and/or field studies;
- reading articles for meaning;
- planning a paper or exposition.

In total, 56 references to concept mapping or equivalent activities were located and reviewed. These were selected from a larger number of references because they were potentially relevant to the application of concept mapping to meaningful learning, as in the present study. A significant proportion of the concept mapping literature, however, relates to its use as an assessment technique. Such literature will only be referenced where it is of direct relevance to the concerns of this study. Here, both concept mapping procedures and the claim that constructing concept maps leads to improved, meaningful learning will be examined, and gaps in the literature identified. In the next chapter, specific research questions will be developed on the basis of this review.

4.3.1 Introducing Concept Maps to Children

When introducing concept maps to children, Novak & Gowin (1984) recommend starting with an explanation of what a concept is. In the case of the youngest children, it is suggested that children learn to differentiate between words that trigger an image in the mind and linking words. In later years the term "concept" is explicitly taught. There are difficulties with this approach, which soon breaks down. One is that it reflects an empiricist conception of concepts, grounded in sense impressions, rather than a social constructionist view, in which the meaning of a construct is how it is used. The ability to form an image is a common, but nonessential, aspect of concept possession, and White (1988) suggests that people's capacity for forming images may vary considerably. Moreover, linking words are themselves concepts, and may therefore cue images. "Makes", "helps" and "blows" have all featured in children's concept maps during the course of this present research, and each of these can conjure up an image for some people. Whether the meaning of "concept" can be taught effectively to children in the earliest years is a question which must at present remain open. But the extent to which this is necessary at all will depend upon the circumstances under which maps are to be constructed. Where children are given a list of concept labels to work with, an accurate notion of "concept" may be something that is unimportant at the outset, but which experience of constructing concept maps would help children to develop.

4.3.2 At What Age Can Pupils Begin to Benefit?

The major proportion of research into concept mapping has been carried out with subjects at high school level and above (aged 12 years to adult). In particular, there is a dearth of systematic research into the effectiveness of the technique at the primary/elementary school level; in a meta-analysis of concept mapping studies, Horton *et al.* (1993) identified only two quantitative studies for this age range, neither in the field of science education. In addition to these, Symington & Novak (1982) developed ways of teaching concept mapping to Australian primary school children, but do not report on the results in terms of pupil's learning. Stice & Alvarez (1986; 1987) also report success in teaching young children to construct hierarchical concept maps, in what appears to have been an action research project. They claim that children so taught showed improved understanding of conceptual relationships, but present no evidence to support the claim. A study by Curry (1984) could be traced only in the form of an abstract, but is of interest because it suggests that measurable gains in learning are possible with elementary school children using concept maps as aids to reading science texts. Compared with a "traditional read-study approach", concept mapping led to "superior" performance on tests of comprehension. Unfortunately, neither the size of the effect nor it's statistical significance is reported in the abstract. This finding from small-scale research would benefit from support from further studies with a similar age group.

4.3.3 Does Ability Matter?

Evidence on this point is equivocal. Stensvold & Wilson (1990) used standardized tests of educational development and tests of comprehension of the domain being learned (chemistry) to examine the possible interaction between various aspects of ability and the effect of a concept mapping treatment. They found that ability had no significant effect on the quality of concept maps produced, and less able mappers performed better on the posttest than similarly able control subjects. They noticed, however, that there was a level of verbal ability above which the treatment seemed to have no effect on performance. Heinze-Fry & Novak (1990) found no statistically significant treatment effect for concept mapping in the domain of college level biology. However, all differences were in favour of the mappers, and the differences were greater for the more able students. They conclude, on the basis of anecdotal evidence only, that lower ability students would benefit over a longer term. Schmid & Telaro (1990) found significant learning effects for low, medium and high reading ability high school biology students using concept mapping. The effect was greatest for low ability students, then high ability, then medium ability. It has been suggested that the introduction of concept mapping may to some extent disrupt learning for pupils with established study techniques (Stensvold & Wilson, *op cit.*; Barenholz & Tamir, 1992). One solution proposed by Barenholz & Tamir (*op cit.*) is to ensure that concept mapping is introduced to pupils early, so it may be that the question of ability is less critical than that of the age at which concept mapping may be introduced successfully.

4.3.4 What Kinds of Learning Gain have been Investigated?

In studies of learning through concept mapping, one type of assessment instrument predominates; objective tests, primarily of multiple-choice format. This is unsurprising, as this is one of the main means of assessment used by science teachers (White, 1988). Multiple-choice tests are highly efficient for measuring certain types of learning outcome. In particular, knowledge of discrete pieces of information is easily assessed by this means, and with a little ingenuity, multiple-choice questions can also be used for measuring understanding. For example, useful multiple-choice questions may be written such that their distractors prompt known common errors or misconceptions (White & Gunstone 1992). However, this assumes that the distractors really do represent typical modes of thinking in the age group concerned, and ensuring this would be a major research project. Typically, a multiple-choice question will give the respondent a choice of four or five possible answers, none of which need correspond exactly with her or his personal beliefs. Also, recognizing an answer and producing an answer make different demands on memory, and so do not necessarily reflect the same aspects of cognitive organization.

Objective tests pose characteristic problems for pretest-posttest designs. In order to measure learning adequately, they need to be of sufficiently low facility to avoid "ceiling" effects (Linn, 1988). This in turn means that they may not afford sufficient discrimination between pupils when used as a pretest (the "floor" effect; *ibid.*). Multiple-choice items are especially prone to this, as subjects' scores may be grouped around the "chance" score which would be expected on the basis of random guessing. Mean pretest scores were not quoted in all the research papers reviewed, making it impossible in these cases to judge the extent to which this might have occurred.

There is a further limitation in the use of multiple-choice questions in this context. If identical tests are used for pre- and posttesting, pupils are alerted, not only to particular questions they have found difficult, but also to possible answers. Any gains measured may, in part, be due to the effect of the pretest in alerting pupils to possible answers to these problem questions. In order to avoid pretest sensitization, Pankratius (1990) used a six group design to assess two levels of concept mapping treatment against two control groups (the two levels of treatment are of interest, and will be discussed later). One of each of the treatment groups and one of the control groups were not pretested. Analysis of variance failed to detect any significant sensitization effect due to the pretest. However, a significant effect[4] was found for both levels of treatment, as reflected in scores on a selection of item bank questions. Although this attests to the effectiveness of concept mapping on attainment, and therefore to its overall positive effects on students' cognitive organization, the nature of the changes to cognitive structure cannot be determined.

As with a number of other studies reviewed, Pankratius (1990) does not specify the exact nature of the assessment instruments used. Even where research papers do describe the type of test, it is uncommon for examples of items or any evidence of validity to be provided. It is therefore, in these cases, impossible to judge the appropriateness of the questions for the purpose of evaluating meaningful (as opposed to rote) learning.

This discussion is intended, not to dismiss altogether the findings of any research making use of objective test items, but to put these into some perspective, and to suggest that they need to be supplemented by other kinds of information, gained from using alternative measures.

In this respect, an important, and often referenced, study is that by Novak, Gowin & Johansen (1983). For this research, 7th and 8th grade US high school students were taught the techniques of concept mapping and "Vee" mapping as part of a programme to improve their learning skills. (Vee mapping is a learning approach described fully by Novak & Gowin, 1984.) The other elements comprising the programme are not specified. Having established that students of this age could master the techniques, Novak *et al.* (*op cit.*) went on to determine the extent to which the programme affected performance on novel problems. After instruction on the gas laws, experimental and control students were given a description of an event

[4] $p < 0.05$

which could be explained by applying knowledge gained from the instruction. Control group students were asked to plan and present an explanation, making use of certain specified construct terms. Experimental group students were asked first to construct a concept map based on the description, and then to write an explanation making use of specified constructs. The explanations written by the experimental group featured, on average, just over twice the number of valid conceptual relationships included by control students[5]. However, this study raises a number of questions.

Firstly, although other assessments of learning are said to have been included in the study, no results are presented for these. This is an unfortunate omission, as it precludes a more complete evaluation of the learning tools. Secondly, it is difficult to believe that the different versions of the assessment task are truly comparable. Control group students were told which *terms* to include, whereas experimental subjects were told to include the *constructs* in the descriptive paragraph, and also *any others they wished.* Thus, while the control group were only required to attend to particular words, the experimental group were directed explicitly to consider, not only the underlying ideas, but also others that may have been relevant. Further to this, it is plausible that the performance of the control subjects could have been improved if they too had been asked to make an initial outline of the important points, and then to rework this into an explanatory paragraph.

Thirdly, as at least two different treatments were introduced simultaneously, it is impossible to determine the effect size for concept mapping alone. Vee mapping is a strategy aimed at helping learners understand the relationship between constructs and the interpretation of events. It is therefore possible that, having been introduced to Vee mapping, the experimental subjects had a clearer grasp of what constitutes a sufficient explanation of an event. Although subjects were not directed to use Vee mapping in answering the task, it might still have had a "background" effect, which it would be impossible to quantify. The use of Vee mapping in conjunction with concept mapping is a feature of other research studies (Lehman, Carter & Kahle, 1985; Bar-Lavie, 1988). The

---

[5] $p < .001$

contribution of these studies to knowledge about the effectiveness of concept mapping is therefore limited.

Barenholz & Tamir (1992) used multiple-choice questions backed up by open-ended questions, concept definition questions and self-report questionnaires, which required students to estimate their own level of understanding of key constructs. The students were learning from a new programme in microbiology for Israeli high schools, developed on Ausubelian principles of progressive differentiation. The programme made comprehensive use of concept mapping in planning sequences of instruction, in formative evaluation of the materials, as assessment tools and as learning assignments for students. The learning of students who did construct concept maps as part of the sequence was compared with students who carried out more conventional assignments. It was found that mappers scored higher on nearly all the assessment instruments, and the differences were statistically significant for about a third of the instruments. In no case was there a significant difference in favour of non-mappers. This work is of interest in that it contributes towards understanding the effects of concept mapping, not only on students' conceptual knowledge, but also on metacognitive awareness. It is therefore doubly unfortunate that the data published do not allow either for an appraisal of the precise kinds of learning gains detected, or for the *accuracy* of metacognitive judgements to be compared for mappers and non-mappers.

Interviews were used as a means of assessment in three of the studies reviewed (Heinze-Fry & Novak, 1990; Merrill, 1987; Wallace & Mintzes, 1990). Of these, the latter two were concerned with the effectiveness of concept mapping as an assessment device. Heinze-Fry & Novak (*op cit.*) used interviews to support objective measures in evaluating the use of concept mapping by college level biology students. They found modest differences in learning, in favour of the experimental group, which did not attain statistical significance.

4.3.5 Learning Effects due to Concept Mapping

Horton *et al.*, (1993) conducted an extensive meta-analysis of concept mapping studies, based on effect sizes. In addition to the studies included in their analysis, a number of others were located in which concept mapping carried out by students was the only experimental treatment, or in which the effect of concept mapping can be separated from that of other aspects of the treatment. Effect sizes were calculated for these, according to the formula:

$$\text{Effect Size} = \frac{\text{Mean Score (Experimental) - Mean Score (Control)}}{\text{Standard Deviation (Control)}}$$

(ref. Borg & Gall, 1989, p.363)

Table 4.1: Effect Sizes in Concept Mapping Studies

Reference Domain; Instrument	Sample n;	 Mean age (Y, M)	Effect size (SD units) (+ve effects in favour of treatment)
Abayomi (1988) Earth science; n/k	156	n/k (Grade 8) [1]	0.15 [1]
Barenholtz & Tamir (1992) Biology; multiple-choice + justification + open-ended + definition questions	180	n/k (Grade 10)	0.00 (multi-choice) 0.13 (justification) [2] 0.14 (open-ended) [2] 0.17 (definition) [2]
	136	n/k (Grade 11)	0.66 (multi-choice) 0.16 (justification) [2] 0.12 (open-ended) [2] 0.55 (definition) [2]
Basili (cited by Horton *et al.*, 1993) Chemistry; n/k	49	n/k (Grade 13) [1]	0.12 [1]
Bodolus (cited by Horton *et al.*, 1993) Marine science; n/k	244	n/k (Grade 9) [1]	0.06 [1]
Heinze-Fry & Novak (1990) [1] Biology; multiple-choice + interviews	37	n/k (Grade 13) [1]	0.52 [1]
Huang (cited by Horton *et al.*, 1993) Chemistry; n/k	129	n/k (Grade 13) [1]	0.21 [1]
Jegede *et al.* (1990) Biology; multiple-choice	50	16,1	2.02 [1]
Martin & Lucy (cited by Horton *et al.*, 1993) Biology; n/k	31	n/k (Grade 9+) [1]	0.48 [1]
Okebukola (1990) Genetics; multiple-choice	138	16,7	1.82
Ecology; multiple-choice	138	16,7	1.67
Okebukola & Jegede (1988) Biology; multiple-choice	145	16,2	1.63 [1]
Pankratius (1990) Physics; item-bank questions	87	17,0	0.21 (low treatment) 0.60 (full treatment)
Schmid & Telaro (1990) Biology; multiple-choice + open-ended questions	43	n/k (Grade 9+)	0.11 [1]
Spaulding (cited by Horton *et al.*, 1993) Chemistry; n/k	44	n/k (Grade 11) [1]	-0.31 [1]
Biology; n/k	107	n/k (Grade 10) [1]	-0.13 [1]
Stensvold & Wilson (1990) Chemistry; n/k	104	n/k (Grade 9)	0.35 [1]

Notes
1 Data published by Horton *et al* (1993)
2 Mean effect size from several subtests

The combined results are presented in Table 4.1. Information in the table reveals some interesting patterns. Firstly, there is a preponderance of studies of concept mapping in biology, followed by chemistry. Secondly, the effect sizes vary enormously. The reported results tend to suggest that concept mapping *can* have quite considerable effects. However, caution is needed in interpreting these results, as some researchers used uncorrected posttest scores, some gain scores, and some corrected scores (used in analysis of covariance). Also, there were differences in approach to concept mapping tasks, with some students' being given lists of concepts to include in their maps, and others choosing their own. Having made these provisos, it does appear that concept mapping almost invariably has a positive effect on achievement, mainly as indicated by multiple-choice tests. However, there is a scarcity of valid evidence supporting the claim that concept mapping improves learners' ability to integrate new information effectively in cognitive structure.

4.3.6 Concept Mapping and Collaborative Learning

It was suggested above that individual concept mapping may aid reflective thought, but that socio-cognitive conflict arising out of collaborative concept mapping may increase the efficacy of the technique by encouraging negotiation of scientific meaning. However, although it has often been claimed that collaborative concept mapping is particularly beneficial (Novak & Gowin, 1984; Stice & Alvarez, 1986; 1987), there have been few attempts to collect evidence for this. Cleare (1983) used concept maps made by student teachers as a focus for discussion and clarification of understanding; some students worked as pairs, while others worked individually on their maps. Resulting concept maps were scored, and analysis of covariance was used to measure learning gains. Cleare reports a difference in favour of the individual condition, but the reported F value of less than one suggests that this difference was not statistically significant.

Okebukola & Jegede (1988) compared the achievement of 15- to 21-year-old Nigerian science students using concept mapping, either in cooperative groups of five members, or individually. Cooperative groups were formed to be of heterogeneous ability. Differences in achievement, as indicated by a multiple-choice test, were found to be significant[6] and in favour of the

[6] $p < 0.01$

cooperative condition. The study also investigated the effect of cognitive preference on achievement, and this was found to interact significantly with leaning mode. These results are difficult to interpret, however, as assignment to the cooperative or individual conditions was determined by individual preference. It cannot, therefore, be assumed that a fair comparison was made. Had the students been randomly assigned to a learning condition, or had the groups been reversed for a second trial, then different results might have been obtained. As the results stand, they cannot support the conclusion that "*making* students work together cooperatively on concept mapping tasks is more likely to improve their performance.....than *making* them work individually" (*op cit.*, p.499, emphasis added).

In the study by Barenholz & Tamir (1992), students were asked to report their attitudes towards aspects of concept mapping. In particular, they were asked whether they liked to discuss their maps with friends. There was a pronounced negative attitude towards this proposition, which suggests that, at least in the particular circumstances of the study, there may have been reluctance to share ideas about concept maps. However, cooperative mapping does not appear to have been emphasized in the study, and the wording of the questionnaire item suggests that "ownership" of the concept maps was individual, rather than collective.

It may be concluded, therefore, that there is a lack of valid evidence that cooperative concept mapping is of more value than individual work. However, Okebukola's & Jegede's work suggests that cooperation may be beneficial. It is also possible that group size is a factor, as Cleare (*op cit.*) did not find a significant difference between pairs and individuals. It is, though, quite possible that group size and the individual/cooperative option may operate differently with different age ranges, and in different social contexts. In particular, the data available may be of little relevance to concept mapping in the primary school age range.

In marked contrast to the research reviewed up to this point, Roth & Roychoudhury (1992; 1993; 1994) used qualitative and interpretive approaches to examine how collaborative concept mapping affected learning for various private high school physics students over a number of years. In the course of the research, discussions arising in concept mapping sessions were recorded and analyzed to identify processes at work. Their findings suggested that concept maps could indeed support the kind of collaborative cognition advocated in this thesis:

> In these discussions, students verbalize tacit knowledge, their own conceptions, and make them available to critique, inspection, discussion, and personal reflection. (Roth & Roychoudhury, 1992, p.549)

In this process, the concept map served as a tool for communication amongst students and between students and their teacher, by way of providing a common referent. This produced sustained discussion, and the students developed together the propositions incorporated in the maps, adding or changing content over a number of turns in the discussion. Roth & Roychoudhury (1993) claim to have identified three major processes in the discourse: ***collaborative construction of propositions, adversarial exchanges*** and the ***formation of temporary alliances***. These, they point out, mirror the kind of interaction associated with communication in scientific communities. However, although generally supportive of learning, these processes did not always result in the challenge of misconceptions. Again, these are findings that apply to older students, and it is not known how they relate to concept mapping as practised by younger children.

4.3.7 Amount and Timing of Concept Mapping

The next variables to be considered are the amount of concept mapping that subjects undertake, and how this is integrated into the learning sequence. Pankratius (1990) posited that learners who constructed concept maps immediately prior to a teaching unit, and then subsequently revised their maps as a result of teaching, would out-perform those who only constructed maps at the end of the teaching unit. The subjects were 87 high school physics students. As stated above, performance was measured in relation to a 30 item objective test, which was matched to the content of the teaching unit.

The study appears to have been carefully constructed to avoid pretest sensitization (see the discussion above). Pankratius quotes a statistically significant difference favouring the concept mapping groups over the control groups[7], and also a significant difference favouring the "high level" treatment over the lower level treatment[8] (*ibid.*, p.321-3).

---

[7] $p < 0.05$

[8] $p < 0.05$

The effect sizes are given in Table 4.1. The substantial effect for the "full" treatment, and the size of the difference between the two treatment methods, suggests that using concept maps as an integral part of the learning sequence is particularly effective, especially when compared to using concept maps simply to summarize learning following instruction. It appears, then, that the construction of an initial, pre-instructional concept map may be an effective way of "laying the foundations" for subsequent learning.

A threat to validity in Pankratius' study is the large loss of subjects over the course of the research (from an initial 145 to 87), some of which appears to have been selective dropout. There is reason to suspect that students in the high treatment level/posttest only group were particularly highly motivated (*ibid.*, p.324). As this group achieved the highest posttest score, the effect size for this group could be overestimated. The result should be accepted with some caution, therefore.

4.4 The Need for Further Research

In this chapter, it is shown that evidence on how concept mapping might contribute to children's learning in science is at the same time voluminous and narrow. A great many previous studies have supported the claim that making concept maps in the context of a scheme of work will lead to improved learning. The main bulk of research has been conducted with pupils in the later stages of compulsory schooling, or often students of beyond school leaving age. Moreover, with the notable exception of the studies carried out by Roth & Roychoudhury (1992; 1993; 1994), almost all the published research has been conducted within an experimental paradigm using a restricted range of measures. These have typically used a pretest/posttest design, with groups assigned to either a treatment (concept mapping) condition or a control condition ("conventional instruction"). Such studies tell us a lot about the extent to which the *average* scores of pupils on objective test items can be increased by engaging in concept mapping. They tell us nothing at all about what processes contribute to these gains. The result of this narrow focus is that little is known about the *manner* in which concept mapping has its effects. The extent to which learning gains are due to a more closely interrelated, and hierarchically organized, personal cognitive structure is, for example, unclear. There is a need for research which investigates the effect of concept mapping on learners' cognitive organization.

Although there is a limited need for further evidence on whether concept mapping is effective in improving test performance in science for a range of age groups, there appears to be a far greater need to understand how to make the best use of concept mapping. Experimental studies typically treat concept mapping as a "black box": so long as there is an observable effect in the appropriate direction, it is not necessary to know how that is brought about. The primary classroom teacher has a different agenda, and is unlikely to adopt any teaching-learning approach on the basis of effect on test scores alone. Teachers need to take on a range of methods that mesh with the overall ethos of the classroom, and that produce effects consistent with a broader range of aims than merely improving learning in a single domain, however important that may be. Many teachers also recognize that a technique or approach that suits one child may not suit another nearly so well. This is something about which focusing on averages tells us nothing at all. Intelligent use of concept mapping and improvement of concept mapping practice are only likely to result from an increased understanding of the processes involved. No amount of information from quantitative research is of any use unless we know how to *interpret* the outcomes: experimental results do *not* speak for themselves. This is a subject upon which the research literature can shed very little illumination, although in the introductory passages to various publications there can be found speculation about how concept mapping might function to improve learning.

In connection with this speculation about how concept mapping might assist learning, it has often been suggested that collaborative mapping should be particularly effective as a learning activity. Yet, as was also shown above, the evidence to corroborate this claim is not secure. Here again, there is a need, not only for more experimental results, but also for a better-elucidated framework of understanding through which those experimental results may be interpreted. Hence further data are needed, both to compare individual with group concept mapping, and to show how concept mapping is actually undertaken by groups of learners.

In the next chapter, these identified gaps in the research literature will be translated into specific questions for the empirical element of this research.

5

INTRODUCTION TO THE EMPIRICAL STUDY

5.1 The Aims of the Research

The present investigation was conceived in order to further the range of information on how concept mapping might contribute to children's learning of science. The focus for the investigation was children in the upper years of the second key stage of primary schooling. In England, this means children aged between nine and eleven years. As was stated in Chapter 1, progress in the "knowledge" programmes of study for science, at this age, is expected to be marked by a developing grasp of certain scientific constructs and associated vocabulary. The *prima facie* case for using concept maps as an aid to increasing children's facility with these ideas was made in Chapter 4. At the time the present research was carried out, there was limited evidence to support this case, and so one aim was to contribute such evidence.

In the preceding chapters, a theoretical basis was developed for using concept maps, a basis that is rooted in a philosophy of learning compatible with that in many primary classrooms. It is against this background that concept mapping will be examined. The research therefore had the following primary aims:

- to develop appropriate methods for using concept mapping with children in the upper years of National Curriculum Key Stage 2;
- to investigate whether concept mapping executed in this way affects children's learning in science;
- to investigate whether collaboration between children when constructing a concept map contributes to any learning effect identified, and if so to identify and describe the processes involved and how these might contribute to learning.

These aims were translated into three research "phases", as described below in 5.4. The latter two phases had specific research questions that evolved over the course of the research, building on findings from the previous phases. These questions were:

Q1: *Does concept mapping, as developed in this study, result in increased integration into cognitive structure of those constructs included in the pupils' concept maps, as compared with pupils learning the same domain who do not engage in concept mapping?*

Q2: *Do the concept maps produced by children in accordance with the principles developed in this study show evidence of an appropriate progression in structure from the pre-topic map to the post-topic map?*

Q3: *Does collaborative or individual concept mapping better promote the development of scientific meanings?*

Q4: *What processes characterize the discussion and production of a concept map by these groups of children?*

Q5: *Is the group production of a concept map best characterized as a constructive or a reconstructive activity?*

Q6: *Does the emerging concept map help to structure the children's activity in a way that encourages the critical sharing of meanings and the emergence of new understandings?*

Q7: *How do the processes at work in the groups relate to the broader views about learning in science developed in the first three chapters?*

5.2 Epistemological Basis

The basis of a research design for approaching the above aims and formulating the above questions will now be discussed. No single methodological "paradigm" or approach will be favoured, and this in itself requires justification. Hence the first area to be discussed will be the broad epistemological commitments underlying this research.

Firstly, it is necessary to establish what is to be taken as "given" in this research. It has already been demonstrated above that an indispensable basis for any stable knowledge must be an assumption of a mind-independent "world" to which we are able to make consistent reference (see, for example, Harré, 1986). This is one aspect of the given. Then, humans construe features of this world, and in so doing they construe those features *as* something, which is to say they ascribe *meaning* to them (see, for example, Sainsbury, 1992). How they do so is at the centre of this study. The world can accommodate an infinity of different descriptions from different perspectives, but will not tolerate just *any* description. Meaning arises out of humans' communications about the world (which includes each other), in the context of their goal-directed activities. So the second aspect of "the given is -so one could say- *forms of life*" (Wittgenstein, 1967, p.226). The system of constructs that humans develop in the course of their activities are a mixed bag consisting of abstract ideas having no tangible referent, but which are useful for summarizing aspects of the world that bear on our experience (constructs such as "niche" in biology and "true score" in test theory) together with genuine but provisional attempts to refer directly to real entities (constructs such as "animal" in biology and "quark" in physics). The balance of these different kinds of construct is most probably dependent on the nature of the form of life.

If we can refer to the world in any number of ways, then there is no single correct or complete meaning system (but again, not just any system will do). Hence we are not concerned with identifying, cataloguing and measuring things and processes in the world *in themselves*. We are interested in how we can classify aspects of the world that are significant in relation to particular purposes (Peshkin, 1993). By way of example, in a controlled experimental study using a pretest and posttest, it is not the scores on the tests *per se* that are the central concern (for all their "objectivity"). It is the *meaning* that can be attributed to those scores that is crucial (Messick, 1989). Even in a paradigmatic example of positivist research, there is, then, an element of interpretation involved. This suggests that traditional ways of carving up epistemological bases for research such as positivist/interpretive, inductive/deductive and quantitative/qualitative are not the most informative. Rather, the emphasis will need to be on the *balance* of these aspects, as determined by the specific research questions. Goetz & LeCompte (1984), for example, suggest that distinctions such as these are better regarded as continua than as dichotomies. Similarly, Hammersley (1991) argues that commonly held dichotomies are better viewed as dimensions, and that adopting a position on one of these dimensions does not necessarily commit one to a particular position on another dimension.

One point at which a distinction may be made is that of the rôle of theory in the research. When there is an established underlying theory, which generates clear and falsifiable hypotheses, then the stance will tend towards deduction/falsification. This tends to be when a theory is reasonably mature (Kuhn's "normal science") and there is a period of stability, not only in the physical world, but in our interpretation of it. In turn, this situation may engender variables that are sufficiently well defined and understood as to make measurement and quantification possible. A model of hypothesis testing as in natural science may then be used, with careful isolation and control of variables. When the theoretical framework is not as well articulated, then development of theory may be the priority. Hence one starts, not with a theory, but with phenomena that one wishes to understand. From here, work commences to gather as wide a range of data as possible, in order to discover categories and patterns from which a theory can be elaborated. This, broadly, is the "grounded theory" approach to research (Strauss & Corbin, 1990), which is towards the inductive end of the deductive/inductive continuum. However, although scarcely acknowledged by Strauss & Corbin, this approach is not (and cannot be) theory-free at the outset. As Constas (1992) has written:

> Contrary to what some have claimed, categories do not simply "emerge" from the data. In actuality, categories are created, and meanings are attributed by researchers who, wittingly or unwittingly, embrace a particular configuration of analytical preferences. (p.254)

Theoretical assumptions and preconceptions underlie each step of the process, and without these it would not be possible to decide on what constitute "data", on appropriate ways of locating or collecting them, or on how to make the first move in the analysis. These are important issues that need to be confronted in developing the research design. At present, though, it is essential to note that it is primarily the scope and centrality of the theory or theories that marks the main distinction between the two broad perspectives described thus far, rather than whether or not a theory exists prior to commencing the research. It is therefore possible that deduction and induction, exploration and falsification may all contribute and complement one another in different stages and aspects of a piece of research.

With the distinctiveness of these two broad traditions thus blurred, it is now apposite to examine an overarching perspective through which their interaction may be conceptualized, and which also unites the two aspects of "the given" acknowledged above. Hermeneutics provides such a perspective (Weinsheimer, 1985; Eger, 1993a; 1993b). Eger (1993a) reinstates in a new form the old metaphor of nature as a "book", which we "read" through the activity of science. This marks a commitment to an independent reality. Unlike its construal in the original metaphor, "reading" is nowadays taken to be a more active process in which the reader brings existing understanding and prejudice to the text in order to interpret it (c.f. Smith, 1971). Here there is a commitment to the constructive nature of human understanding. Rather than impediments to achieving scientific objectivity, preconceptions are regarded in hermeneutics as an essential prerequisite to advancing understanding. From such preconceptions, we make a prejudgement about the object of our study, which we then apply in an attempt to come up with a plausible reading. The way the object of study reacts to this probing may reveal anomalies, which require an adjustment of our preconceptions, and their reapplication to create a second approximation. This proceeds, iteratively, until sufficient anomalies are accounted for and a coherent interpretation is made. Ultimately, the basis for what we can prove lies in what we know but *cannot* prove (Weinsheimer, *op cit.*). But although we start from our prejudices (which are related to our purposes), it is the continual process of offering up our partial interpretations to the original book of nature (independent reality) that creates the possibility of convergence on a valid account.

The result of this "reading" of nature is the production of another "book", the book of science, which is essentially an account of how science has interpreted the book of nature. But this second book also requires interpretation (and sometimes reinterpretation) on the part of all those who are to work in the field, who are to continue writing the book. "Initially, the scientist is always a student of symbols created by others." (Eger, 1993b, p.306). One result of conceiving science in terms of this "cascade of interpretations" is the deconstruction of a sharp subject/object distinction. We can shift our attention from more to less objective positions, but at each stage there is an object to interpret and an interpreter. Eger (1993b) draws the analogy with a play, in which interpretation moves from the script (a core that structures possible interpretation) through interpretation by the director and then the various actors and designers in a performance, to the interpretation by each member of an audience (and this will be different for each production and each performance). Neither the object nor the subject's cognitive structures is privileged. Both are significant.

In support of this, Eger (drawing on Polanyi) shows how the scientist is "embodied" in a set of equipment and cognitive tools, like theories and language, which extend her or his perceptual reach. He illustrates this with the case of an astronaut who puts on special suits to conduct various kinds of exploration. To look upon the suit as merely a means to do investigations overlooks the theoretical basis for its existence:

> At the start of the astronautical project, the suit itself had to be the focus of concern, the object: but as that problem was "solved", and the astronaut "entered in", the suit was joined to his body, became more or less peripheral to awareness ... and could be viewed thereafter as part of the subject. What happened is that the subject/object cut shifted in the course of the enterprise. (Eger, 1993b, p.308)

Perhaps a better example would be the development of electromagnetic understanding which ultimately enabled the production of video display screens that in turn provide instrument readings for innumerable experiments. The original theories became part of the periphery within which later work is conducted. Eger takes this movement of the subject/object boundary to be a characteristic feature of hermeneutics. It is also highly relevant to the question of research methodology in social science, for it enables us to characterize ("reinterpret") such distinctions as positivist/interpretive and quantitative/qualitative in terms of a shift of the focus of attention. Someone who conducts an educational experiment starts with a theory which is sufficiently objectified for it to be operationalized in terms of specific consequences. The researcher need not focus at that time on such subsidiary components as the binomial theorem (indeed, as with the space suit, need not understand it), yet they form part of the theory system, the tradition, upon which the approach depends. In the case of grounded theory, there is no less a tradition of pre-existing theories, but they are likely to be less formalized, and the researcher may not even be able to express them. The aim is to develop and objectify a theory (which ultimately might form the basis of an experimental manipulation). None of this should be taken to imply that the often tacit assumptions underlying a particular methodology need not be questioned, either in general or piecemeal, in relation to a proposed use. Far from it. The point is that different methods need not be seen as incompatible with one another, or with a unified epistemological view.

In hermeneutics, then, we have a way to conceptualize the knowledge-garnering processes of science in general (and, *a fortiori*, of social science in particular) that is independent of particular methodological positions. Prior conceptions are central to the collection and valid interpretation of data, and consequently for the generation of knowledge, on this view. These prior conceptions come into play, for example, in a researcher's selection of data gathering technologies, in the appropriation of a well-elaborated theory to be tested, or in a researcher's everyday understanding of the language used by respondents. Within this overarching epistemological position, we can locate different methodological perspectives in different stages of the "cascade of interpretations", and therefore choose among them according to the specific purposes of the research. We also have a framework within which the assumptions behind these methods (the realm of preconceptions or "tradition") can be examined.

The present research forms part of a continuum of knowledge acquisition that in its entirety is a hermeneutic process, and it is on this epistemological basis that the methodologies are selected. Because there are several research questions, the appropriate methodology for each will differ, according to the nature and scope of relevant existing theory. The specific methods adopted will be discussed in detail at relevant points in the following chapters.

5.3 The Research Context

The present research focuses on a classroom-based learning aid. The first decision to be made was *where* that research should be conducted. The choice here was between a "clinical" context, in which as many extrinsic variables as possible could be controlled, and a "natural" context that resembled as closely as possible the conditions under which the approach would be used in a classroom. The solution to this question draws on the perspective provided by symbolic interactionism (see Jacob, 1987). From this perspective, humans are said to "act towards things on the basis of the meanings those objects have for them" (*ibid.*, p.27), meanings which are formed through social interaction in regard to those things. The meaning a pupil attributes to a situation is therefore of central importance in achieving valid research findings. Since the social context of clinical research is likely to be somewhat different to that of the classroom, where the rules, social relationships and activity structures are already familiar to the child, results that generalize beyond the specific situation of the research are more likely to be generated in the latter context. In the unfamiliar setting, the child is having to make sense of the situation, including what responses the researcher values, as well as of the task. Consequently, she or he may perform rather differently under these conditions. Hammersley (1991) refers to this as "reactivity" to the research setting, which, depending on its degree, is a potential threat to "ecological generalizability"[9], that is, the extent to which the findings can be generalized to another setting. To maximize ecological generalizability, the research needed to take place in real classrooms, and needed to blend as far as possible with normal classroom activity. This called for a research approach that was as "naturalistic" as possible, and which minimized the use of obtrusive data collection methods. Concept mapping needed to be carried out for genuine classroom purposes related to the work the children were actually doing in science, and to be evaluated through criteria related closely to that work wherever this could be achieved. However there were constraints on the extent to which these ideals could be attained.

[9] Hammersley (*op cit.*) uses the term "ecological validity", a term which Messick (1989) considers unsatisfactory.

Running alongside the requirement for naturalistic research was a moral imperative that the "subjects" involved, the children, should be as informed as possible about the purposes of the research. This stemmed from a desire to have the children as consenting participants, and therefore of allowing them to raise objections. The children needed, in any case, an explanation for the presence of the researcher and his data collection apparatus in their classroom. Clearly there is a tension between making children aware of the purposes of the research and avoiding distortion of the results. This tension could, however, be reduced by phrasing the aims of the research in terms of "finding out about" and "seeing if there are any differences", rather than stating expectations about the findings.

The second constraint occurred in those situations where it was necessary to make direct comparisons between learning approaches, since this demanded a high degree of control over factors which might affect the variables to be compared. This inevitably injected an element of artificiality into the classroom. The approach taken was to minimize this artificiality as much as possible by identifying classroom contexts in which conditions favourable to the research activities already existed. The different parts of the research may be placed at slightly different points along the artificial/natural continuum.

5.4 Dependability of the Research

Dependability of research outcomes is traditionally discussed with reference to the constructs of *validity* and *reliability*.

5.4.1 Validity

Messick (1989), although writing in the field of educational measurement, has nevertheless provided an analysis of the construct "validity" that is philosophically sufficiently well-grounded as to be generalizable to all forms of data gathering. Central is the idea that validity concerns the *meaning* that can be attributed to the data collected. Inferences made from data require *justification*, and this is provided by marshalling evidence of various kinds to sustain an argument. In terms of the present study, this translates into the questions:

- are the data gathered the ones appropriate to the research questions?

- are the inferences based on those data defensible (or are there alternative inferences that could be drawn)?
- how may the inferences based on those data be generalized?

To address the first of these questions, the theoretical basis for the collection of data will be specified. The second question is more complex. A first step is accomplished through creating and documenting theoretically justified analysis procedures. The next consists in ensuring that these are enacted consistently (that is, reliably). Also, where more than one analysis is drawn on, these should converge on the same result. To discount alternative inferences, it is necessary to eliminate sources of error, such as uncontrolled factors that might affect the performance of groups. These aspects of validity argument are essentially *internal* to the study, that is, they ensure the inferences made about the specific groups involved are appropriate. To demonstrate how validity was ensured in these respects, the following sections and chapters carry a full description of the procedures implemented.

The last of these questions, in contrast, concerns whether the results have any *external* applicability, to other times and other groups of children and teachers. One aspect of the question is directed at the sample used in the research, and of what population this might be representative. It points to what many would regard as the Achilles' heel of qualitative research. Because of the depth of analysis required, the number of cases tends to be low, and mathematical theories of probability cannot be relied upon to answer questions about generalizability. Yet even in instances where statistical significance can be determined, we cannot be certain about generalizability. We can only say how *likely* it is that the results were not obtained by a fluke of the sample, and only then if the sample satisfies the randomness condition (which, given that one is usually dependent on voluntary participation, is hardly ever guaranteed in educational research). Some qualitative methodologists might argue that generalization is not their aim, and that what is wanted is a firmly contextualized description, a description that is only valid for the situation in which it was developed. But such a response would miss the point that one always has a reason for doing the research, and therefore *some* kind of generalization in mind. A description that is only valid for one instant in time and in one situation is of no value. Even ethnography, perhaps the epitome of situated research, has its origins in the quest to understand our own society better (Goetz & LeCompte, 1984).

Valid generalizations are therefore dependent on avoiding the situation where the sample is for some reason untypical. Under these circumstances, statistical tests may be regarded as a guide to generalizability. However, this is still in essence a probabilistic view, and for political reasons insufficient. If the position is taken that consumers of research (especially teachers) should not be mere recipients of the findings, then it is necessary for them to be enabled to make the judgement "could the findings apply in my individual situation; a context different to that in which they were generated?". Hence as well as ensuring that the sample is as far as possible suited to the purpose of the research, we shall also follow the direction indicated by Peshkin (1993). Drawing on Wehlage and others, he shows that generalizations may be located "in the relationship between text and reader" (p.26). On this view, the onus is on the author to provide the reader with information to judge the extent to which findings are likely to apply in a certain context.

5.4.2 Reliability

Reliability is the term used to denote consistency of measurement (Traub, 1994). A data collection procedure that produced substantially different results on different occasions, or with a different user, would be of little use. Hence it is necessary to ensure that the procedures employed are capable of delivering equivalent results across the conditions of their use. Checks on reliability are based on comparing results obtained under different conditions. The exact methods employed depend on the nature of the data and the assumptions that may be made about them.

Classical reliability theory is applicable where numerical scores measured on an interval or ratio scale are involved (Traub, 1994). Here one is generally concerned with the magnitude of measurement error, and consequently with the correlation between the score obtained and the notional "true" score. To avoid ontological problems, this true score is best conceptualized as the average of the scores that would be obtained under the range of possible conditions of use. It is estimated by finding the correlation between the scores obtained under contrasting conditions.

Where categorical data are concerned, it is the level of agreement that is of interest, and it does not make sense to talk of the magnitude of an error. Consequently, the method used compares the decisions reached under contrasting conditions (see Cohen, 1960). The results of reliability checks used in the present analyses are presented at appropriate points in the discussion of the research design.

5.5 The Overall Plan

The research was timetabled to take place over three phases.

1. In the first phase, ways of introducing and using concept maps effectively with primary-age children were identified. The procedures were then tried out with small groups so as to confirm and, as necessary, improve them before using them in the main research. During this preliminary phase, research questions were clarified, to guide subsequent phases.

2. In the second phase, the main focus was on obtaining quantitative data on the effects of concept mapping on the structure of children's scientific knowledge. This phase was intended to establish whether concept mapping has any effect on learning in science at the age range of interest, thereby warranting a search for *how* it might support learning. A study using an experimental design was set up to make a preliminary assessment of the significance of any learning gains identified.

3. The third phase was reactive to the outcomes of the second phase. Part of the research consisted of a substudy to compare the effects of individual and collaborative modes of concept mapping. Simultaneously, the main focus of the phase was on gathering qualitative data in relation to collaborative concept mapping that would enable the nature of the processes involved in constructing the maps to be elucidated.

The remainder of the thesis is structured around these three research phases. Phase one, which lays the foundation for the rest of the research, is described in the last part of the present chapter. Chapter 6 concerns phase two of the research; its rationale, its methods, its implementation and its findings. In Chapter 7, the rationale for phase three, the research design and its implementation are described. In Chapter 8, the implications of the overall methodology will be developed in terms of an analysis system for the data gathered in phase three. Chapter 9 presents the findings arising from the analyses carried out. Chapter 10 is the final chapter, and in this the findings from both phase two and phase three will be integrated into a discussion of the relevance of this thesis to the wider realm of research into concept mapping and learning in science.

5.6 The First Research Phase

This phase consisted of developing procedures to be used for introducing and using concept mapping in the classroom. They were derived as far as possible from existing literature, in order to preserve continuity of research evidence, and so a major portion in this research phase consisted of synthesising previous recommendations. However, these existing procedures are flexible, and allow for a number of different approaches. The choices made in specifying an approach and the reasons for these will now be detailed. In this phase, draft procedures for teaching concept mapping were written, trialled and revised in the light of children's reactions.

5.6.1 Who Should Choose the Terms to be Mapped?

Learners may themselves select the constructs to be included in a map, or they may be given a list of terms to include, perhaps with the freedom to add further relevant constructs of their choice (Stewart, Van Kirk & Rowell, 1979; Symington & Novak, 1982; Novak & Gowin, 1984; Ault, 1985; Cohen, 1987). The choice made would be determined by the purpose of the concept map, and the specific context in which it is used. If, for example, the map were intended to help the learner make sense of a specific passage in a book (a purpose identified in some of the research literature on concept mapping), then selecting from the passage the most important ideas for inclusion would itself be a valuable learning task. If, on the other hand, the purpose were to have learners reflect on what they know about a particular set of ideas, then it would be necessary to provide them with a list of those core ideas.

For the present study, concept mapping was not based entirely on a specific passage in a textbook. Rather, it was used in the context of a whole scheme of work devised by the children's regular teacher (and therefore largely beyond the control of the researcher) which included diverse elements such as practical work and class discussion. In order to help the children reflect on this range of experience, and to draw from it the important ideas and their relationships, it was decided to give the children a list of the main constructs in the topic. However, to encourage the children to reflect on other ideas which might be related to those listed, they were invited to add whatever further terms they chose. In order to help the children reflect on the relationships between the constructs specified, and in particular to reflect on their hierarchical organization, the lists of terms were supplied in the form of movable labels, which the children could arrange and rearrange as necessary before finalizing their maps (Novak & Gowin, *op cit.*).

5.6.2 When Should Concept Maps be Constructed?

It was suggested earlier that one function of concept mapping is that of helping learners organise their own thinking about a subject. Partly, this would be a process of reviewing what is known. It would also be, in part, a process of elaborating what is known: clarifying the links between relevant ideas and events, and making what is known tacitly available for public scrutiny. The formal requirement to construct a coherent map may expose uncertainties, inconsistencies or gaps in knowledge, revealing where further learning is needed. This suggests that concept mapping would be a valuable way of "preparing the ground" for new learning, in the sense of ensuring that prior knowledge is ready to receive the new ideas to be integrated. Ausubel (*et al.*, 1978) describes such relevant prior knowledge as "anchoring" new learning in cognitive structure. It was therefore decided that pupils should construct an initial concept map at the very beginning of a teaching unit. Having constructed this initial map, it would then be available for review at appropriate stages in the teaching sequence in order to see how each new learning step related to what was believed previously. A fresh map, constructed at the end of the teaching unit, would help the pupils consolidate their learning from the topic as a whole.

5.6.3 Group Composition and Procedures

It was an aim of the study to examine the processes involved in constructing a concept map collaboratively. To encourage discussion, the children were asked to come to a consensus within the group on the final arrangement of the concepts and propositions in their maps.

The optimum size and composition for a concept mapping group needs to be considered. Bulman (1985) suggests that discussion groups should be comprised of between two and four pupils, and these group sizes were used successfully by Barnes (1976; Barnes & Todd, 1977). Although it would be possible to experiment with different numbers and mixes of pupils, the previous experience of the class is likely to have some considerable bearing on this question, and there were therefore grounds for adopting a more pragmatic and flexible approach in this research.

Webb (1982a) reviews the results from three of her studies investigating the effects of group ability composition on the interactive behaviour of individuals working cooperatively in the groups. The results do not yield a firm conclusion, as not all ability combinations were investigated. However, it appears that medium-ability students interacted most effectively in either uniform-ability groups, or in groups where they worked with either higher- or lower-ability students, but not both. Homogeneous high- or low-ability groups were the least successful. Successful interaction in Webb's studies was defined in terms of interactive behaviours that correlated positively with achievement (such as helping, or giving explanations to, others in the group). The findings suggest that cooperative groups are most likely to be effective if there is a mix of abilities which avoids extreme differences. However, this requirement also needs to be tempered against factors in the social ecology of the particular classroom in question. With only a small number of children in the classes involved, it might well prove impossible to achieve theoretically ideal combinations.

In addition to "ability", prior knowledge may be relevant. Johnson & Howe (1978) matched pairs of children in terms of their differing (Piagetian) conceptual development, and found this approach effective in generating cognitive conflict. However, the transfer of this clinical design to the classroom is problematic, as it presupposes an adequate way of detecting conflicting beliefs, and also assumes that there is a convenient balance of these across the class. It would probably be impossible to combine all of the relevant social and cognitive factors when choosing ideal groups in a natural setting.

On these grounds, then, it was decided not to take a strong line on either the size or composition of the group concerned. For the third phase of the research in particular, it was important to preserve natural working conditions of the classes concerned as far as possible. Hence the first priority was to work with groups as they were constituted by the regular class teacher. In classes where the children were used to working in cooperative groups, the teacher could be expected to have determined combinations of personality that worked appropriately well together.

5.6.4 Introducing Concept Maps to the Children

Exploratory work based on the above principles was carried out with a Year 5 class, initially with a group of six children, and then with a whole class of 36. This enabled a simple teaching routine to be developed, trialled, and revised prior to the next phase of the study. The refined procedure is reproduced in Appendix A in the form of a guide for teachers.

In keeping with the principle (developed in the early chapters of this thesis) that learners make sense of new information by relating it to existing ideas, concept maps were introduced to the children in terms of an established referent: a road map. The parallel was drawn between a road map (which shows how places are connected), and a concept map (which shows how ideas are connected).

It was found that the children needed experience of constructing at least two maps of a relatively straightforward domain, and with guidance, before they were ready to attempt a more demanding map independently. (Baird & Mitchell, 1987, similarly report a small number of practice sessions to be desirable in learning to make concept maps.) The children's first map needed to be produced according to a well structured sequence of steps which illustrated the importance of hierarchy in a good concept map. This was achieved by starting with basic level constructs representing observable entities, and then adding higher level constructs to integrate the basic level ideas. The importance of this step and of the resulting hierarchical arrangement of constructs were emphasized. The first map was constructed as a whole-class activity.

For the first demonstration map, and for each subsequent practice map they made, the pupils were provided with a strip of paper on which the construct terms were printed. This was found to be invaluable (a point reinforced by Roth & Roychoudhury, 1993). On one trial visit, the children were asked to construct a map without the use of printed strips. Their maps proved to be very time consuming to produce, and of poor quality. As well as the time taken to write out the list of words and transfer them to the map, the lack of opportunity for rearranging the terms on the page to obtain optimal positioning was found to be a substantial limitation. For the purposes of the second phase of the study, it was considered helpful to give each child working in a group a set of the construct terms, and to construct an individual copy of the concept map, so as to monitor individual involvement in the task. However, it was stressed that the task was to be a group activity, and that all decisions regarding the content of the map should be made collaboratively.

For the first map, the initial connections to be made were demonstrated, and the children then glued their labels in place and wrote in the linking words. Subsequent steps were discussed prior to being finalized. The set of constructs was chosen specifically with a view to illustrating the hierarchical nature of concept maps, and also how further terms, not included in the original list, could be added.

A second strip of construct terms was then given to each of the children, and they were asked to construct a map of their own based on these words. The set of terms was chosen so as to invite a hierarchical arrangement, and was selected from a science topic recently studied by the children.

The children were directed to start off their second map by identifying which constructs in the list were the most "important", and which the least. They were then allowed more freedom to construct their individual maps, whilst their progress was observed to see whether they were applying the principles demonstrated to them. For this second map, the children were given a prompt card which summarized the steps in the construction of a concept map. The card was retained by the children for future use. In a plenary at the end of the session, ideas for the arrangement of concepts in this second map were shared, and the acceptability of different, idiosyncratic, arrangements of concepts was emphasized.

The concept maps produced in these early exploratory trials enabled the success of the procedures to be gauged. A useful yardstick in this evaluation was the presence or absence of linear chains of propositions with no clear hierarchical relationship or cross-linked structure. Since the domains to be mapped were expected to be familiar to the children, lack of a clear structure in the maps produced would tend to indicate that either the purpose or the procedure for creating them was not understood. Such linear chains did not result from the approach as developed here; instead, the maps were hierarchically structured. It was therefore concluded that the instructions were sufficiently clear, and that there was enough practice given, for the children to grasp the principles of concept mapping. As the maps produced by the individual children in each collaborative group were similar in terms of their structure and the words used to link the terms, it was also concluded that there was adequate opportunity to share ideas, although it was recognised that this conclusion was less likely to be generalizable to other settings, where pupils' experience of collaborative working would be different.

The refined procedures described here were therefore used as the basis for the second phase of the research, reported next in Chapter 6. Copies of the procedures were given to teachers participating in the third phase, and in one case was used by the teacher as an approach to introducing concept mapping to her class, which she found helpful in comparison with her previous attempts to introduce concept maps.

6

RESEARCH DESIGN AND RESULTS: PHASE TWO

6.1 Method

This second phase of the research was designed to draw and build upon the first phase. The aim was to test out the proposition that concept mapping carried out in a science domain helps learners integrate scientific constructs within a coherent cognitive structure. The assumption is therefore that the addition of a "treatment" can bring about a significant and worthwhile change in learning outcomes. The test of this assumption entailed the setting up of an experimental situation in which concept mapping could be compared with a more conventional teaching-learning approach, and the adoption of suitable means of assessing and evaluating the effects of concept mapping.

6.1.1 Samples

A true random sample of the population (Year 6 pupils) was not practical for the present research. This was so, not only as a result of limited resources, but also because this would preclude control over a range of extraneous variables. It would have been difficult (and inconvenient to the children's teachers) to insist that selected individuals be taught to specific programmes which might not have related to work being undertaken by others in the class. It would have been practically impossible to ensure that children taught to the same programme by different teachers had equivalent experiences.

Two possibilities remained, both dependent on the samples' being based upon whole classes rather than individuals. Firstly, single classes could be divided, with pupils assigned to experimental and control conditions within the class. Secondly, situations could be identified where classes were taught science in parallel by a single teacher. The latter was deemed to be preferable, on the grounds that it would be less likely to prove organizationally disruptive, given the substantial nature of the "treatment", and would in favourable circumstances result in larger samples.

A number of schools were approached, where it was known that classes were taught science in parallel. One, a private school taking children in the age range 5-13 years, agreed to participate. It was on a modern site in a rural area, but close to a number of commuter-belt towns. Teaching in the upper part of the school was largely by subject specialists, and followed the public schools Common Entrance syllabus, leading to examination at age 13. The school curriculum was also aligned with the National Curriculum for state schools. Science learning was centred on practical investigations, in combination with more formal class teaching.

There were two classes in Year 6, one of 11 pupils (5 male, 6 female) and one of 13 (6 male, 7 female). The age range of pupils in the former group was 9 years 2 months to 10 years 11 months, with a mean of 10 years 5 months. The age range of pupils in the latter group was 9 years 6 months to 11 years 0 months, with a mean of 10 years 4 months. All the pupils came from higher socio-economic group families, spoke English as their home language, and were of average general ability or above.

One disadvantage of this situation was that pupils within the school were not assigned to teaching groups at random, but were grouped according to ability in English. The teacher expressed the opinion that this was not necessarily reflected in differences of ability in learning science, but nevertheless, it could not be assumed that the two groups were entirely equivalent. However, it was also the case that no single measure was used in grouping the children that could be used to control for group differences. This determined both the design of the experiment, and also the kind of statistical analysis deemed appropriate.

6.1.2 Experimental Design

The research design must be termed quasi-experimental, since random assignment of subjects was not carried out (Cohen & Manion, 1980; Borg & Gall, 1989). A pretest-posttest, non-equivalent control group arrangement was employed (*ibid.*), using intact classes as experimental and control groups. For the purpose of constructing concept maps, the children were put into collaborative groups of two or three, in consultation with their teacher.

In order to help reduce threats to validity resulting from the assignment of pupils to teaching groups by ability, it was decided to carry out two experimental runs. In the first of these, the higher ability group would be the experimental group with the lower ability group as the control. For the second run, the assignments would be reversed. For both runs it would be necessary to employ statistical analyses with which to take account of differing characteristics of the two groups.

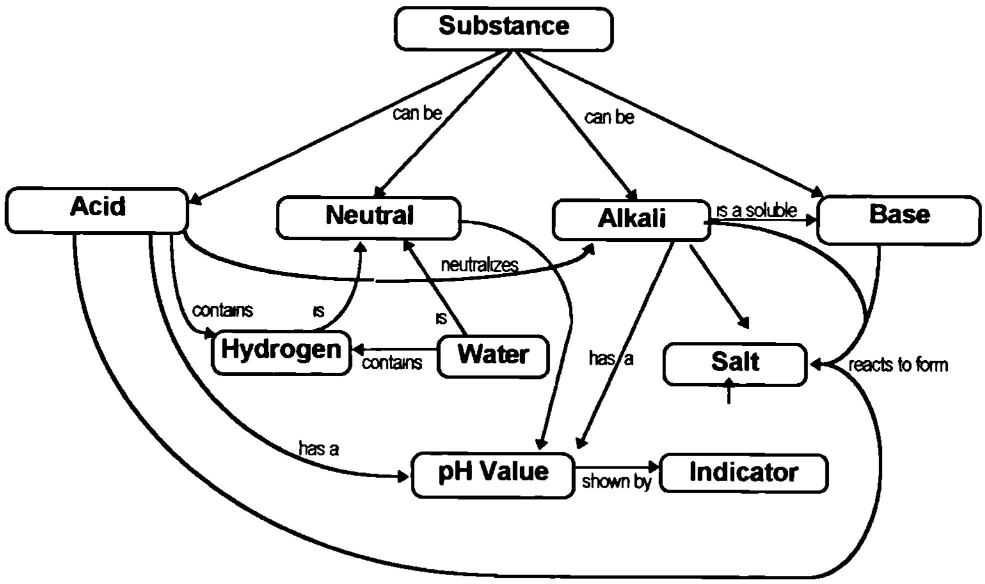
Unfortunately, the children's teacher declined to participate in the second run, and as a consequence, data were only collected for the first experimental run.

Science topics

The science topics that children would be studying during the year were determined beforehand by the school science syllabus, and it remained only to select two topics that would prove amenable to concept mapping. These were chosen, in consultation with the children's science teacher, to meet the requirements that they should feature a number of distinct but interrelated constructs, with an element of hierarchy in their theoretical structure.

The topic used for the first experimental run was entitled "acids and alkalis". The teacher's scheme of work for this topic was based on the chapter of the same name in Harris & Ferguson (1979), a textbook used in the school. This chapter was analyzed to provide the set of construct relationships which would be the focus of the children's concept maps. The main scientific constructs and the propositions relating them were identified from the text. These are presented in the form of a template concept map in Figure 6.1.

Figure 6.1: A Map of the Topic "Acids and Alkalis"



6.1.3 Instruments

In keeping with the epistemological position described in Chapter 5, this experimental study was to be based mainly on established instruments and methods. There are several different techniques that have been used in research to elicit information on cognitive structure. Some of these are described in Archenhold *et al.* (1980), Novak & Gowin (1984), Preece (1976), Shavelson & Stanton (1975) and White & Gunstone (1992). The methods differ in the extent of cognitive structure explored, and the precision with which this is achieved (West, 1980).

Clinical interviews are well respected as a method of probing cognitive structure, but are most effective when focused on a small set of closely related ideas (Novak & Gowin, *op cit.*; White, 1988). The range of concepts to be examined, the human resources available and the time windows imposed for the present study precluded the use of interview methods. Concept maps have also been proposed as a means of examining cognitive structure, but their use in the experimental condition precluded their adoption for the present research.

In view of the relative immaturity of the pupils, the present research required a method of examining aspects of cognitive structure that did not depend greatly on problem-solving skills or other abilities peripheral to the focus of the study. Such a methodology needed therefore to make minimal information processing demands on the children. One technique appears to meet this requirement: the Word Association Test (Shavelson, 1972; 1974; Shavelson & Stanton, 1975. Also, Preece, 1976; Cochaud & Thompson, 1980; Moreira & Santos, 1981; Isa & Maskill, 1982; Fisher, 1986; Cachapuz & Maskill, 1989; White & Gunstone, 1992).

Word Association Tests

The Word Association Test was introduced by Francis Galton and subsequently has been used widely as a probe of semantic relations (Miller, 1991). Although there are several variations on the theme, in all versions the respondent is presented with a stimulus term, and responds by providing (normally in writing) a number of response terms. These responses are considered to reveal something of the meaning that the respondent attaches to the stimulus term. Usually the time allowed for response generation is limited. There may be any number of stimulus terms in a complete "test".

As an example, a hypothetical set of responses to two terms, "black" and "white", might appear as follows.

black	white
white	black
grey	paper
colour	grey
coal	
dark	

In a typical example of a Word Association Test, Shavelson (1972) presented students with fourteen stimulus words from the domain of Newtonian mechanics. One page of the test was devoted to each term. On each page, thirty lines were printed in two columns, and next to each line was printed the same stimulus word. Students were instructed to "write as many words as they could think of" on the lines provided, in response to each stimulus (p.227).

Shavelson (1972, 1974) based his interpretation of the Word Association Test data on a "spreading activation" model of semantic processing (Collins & Quillian, 1969). The model is based on a quasi-neural metaphor, in which it is assumed that the search for connections to a term radiates outwards from the stimulus concept to increasingly distant links, and that the context of the search determines which links are given priority. Thus it can be seen that the context in which the test is conducted is critical to the results obtained. According to Shavelson (1972), the order of retrieval "reflects at least a significant part of the structure within and between concepts" (p. 227). The organization of concepts in memory is what Shavelson refers to as cognitive structure (*ibid.*, p.226).

Variations of the test impose different constraints on the subjects' responses. For example, respondents may be asked to "think like a physicist" (Shavelson, 1974), or they may be asked to respond to physical quantities only with other physical quantities (Preece, 1976). In one interesting version, the context for the Word Association Test was set by having participants solve two chemical equilibrium problems, and immediately afterwards asking them to write responses to terms for key constructs relevant to the problems (Cachapuz & Maskill, 1989).

Three main types of information provided by the Word Association Test are discussed in the relevant literature. Firstly, the *number* of responses to a concept is sometimes used to represent the meaningfulness of the concept, that is, the number of other ideas that are connected with it in memory (Shavelson, 1972).

Secondly, the *order* of the responses may be taken to represent the strength of the association between the stimulus concept and those retrieved. The first responses are assumed to be those most closely related with the stimulus and those that contribute most to the meaning of the concept (Garskof & Houston, 1963). As Collins & Loftus (1975) have observed, humans can generate virtually unlimited information about a familiar concept, but the information generated becomes increasingly marginal to the meaning of the term. Placing a time limit on the test ensures that the connections elicited carry a substantial proportion of its meaning.

Thirdly, the *content* of the responses provides qualitative information on what meaning the respondent attaches to the stimulus concept. Fisher (1986), for example, classified the content of response lists to key construct terms in biology into categories that reflected the type of relationship between the stimulus and response terms (such as subordinate-superordinate relationships, descriptive terms, synonyms and definitions). Doing this across a range of ages of subject enabled her to draw general conclusions about the development of biological concepts. Of course, this kind of interpretation will always be somewhat subjective, and will only be valid to the extent that the respondent and the observer hold compatible meanings for the response terms. It is a weakness of Word Association Test methodology that it provides information on *which* ideas are related, but not on *how* they are related. For example, if a subject responds to the stimulus term "acid" with the stimulus term "alkali", it is not clear whether this is because he or she knows that these are substances that are similar in some way, or whether he or she has simply heard the two words used together on a number of occasions (as in the title of a chapter in a text book).

Some have found in this grounds for rejection of the method entirely (Stewart, 1979; 1980). Stewart's argument is that as associations *may* have arisen simply out of temporal contiguity, a propositional relationship between the associated terms cannot be assumed. While this is fair criticism, he seems to overstate the case, apparently implying that because a propositional relationship cannot be assumed then it does not exist. Even if this were the case, it may be countered that the fact that a respondent makes a connection means that it is part of the context of use of the stimulus concept for that person, and may therefore provide a link to further ideas. Sutton's (1980) criticism is more cautious; that there are large numbers of *possible* connections between some pairs of terms, and the nature of the actual connection made is not known. He emphasizes in particular the responsiveness of cognitive structure to context, and points out that different configurations of concepts may be brought into use under different circumstances. He concludes that Word Association Tests are of little use in revealing the flexibility of cognitive organization, but may be useful for revealing its more static aspects. Provided that these limitations are realized, then they need not be fatal to the methodology. Word Association Tests do allow certain *aspects* of cognitive structure to be inferred; in essence, the Word Association Test is probing the connections to a particular idea within a specific context.

Gunstone (1980) modified the test procedure by requiring respondents to write a sentence for each connection made showing the nature of the relationship. Although this would go a considerable way towards remedying the weakness described above, where the number of stimulus words is large (say ten terms, each generating up to ten responses), the time taken to administer such an assessment would become impractical. The modification was therefore not considered appropriate for the present study.

Word Association Tests in relation to the present study

In several pieces of research (Shavelson, 1972; Shavelson & Stanton, 1975; Preece, 1976; Moreira & Santos, 1981), the content of response lists and the order of the terms in the lists were used to calculate a metric (akin to a correlation coefficient) describing the strength of the connection between pairs of stimulus terms. Essentially, two terms were considered to be related to the extent that they generated similar responses in a similar order, although the order can receive zero weighting if deemed appropriate. The metric used in these studies was the relatedness coefficient, originated by Garskof & Houston (1963), and this has generally been used to derive a structure to the terms listed by such techniques as similarity matrices (Shavelson, 1972) and hierarchical cluster analysis (Moreira & Santos, 1981). In each of these cases, concurrent validity evidence is provided by the authors that supports use of the relatedness coefficient to make interpretations of word association data in relation to significant features of the cognitive structure of the respondents. Further, the cognitive structure thus derived has been shown to change with learning, and to come increasingly to resemble the structure of the content learned by the subjects. It is also related to their capacity to make use of that learning, as revealed in objective tests (predictive validity evidence). This collection of evidence supports the inference that data obtained from Word Association Tests are interpretable in terms of cognitive structure. However, valid use of the relatedness coefficient appears to be based on assumptions that could not hold in the present case, and these assumptions must now be outlined.

The terms used in each of the studies cited were of a similar kind: they were all (or primarily) physical quantities. This situation results in a fairly uniform and symmetrical set of relationships. Where relationships are of differing kinds, then the order in which responses are generated would also be expected to differ. There might well be cases where the meaning of one term is central to that of another, but the meaning of the second term is quite marginal to that of the first, as in the following hypothetical instance.

<u>animal</u>	<u>aardvark</u>
mammal	animal
dog	anteater
cat	Africa
living	

Although it is quite central to the meaning of "aardvark" that it is an animal, very little of the meaning of the term "animal" is carried by "aardvark". Such asymmetry is not recognized by the relatedness coefficient, which is non-directional. Hence, studies based on relatedness metrics are useful in providing validity evidence concerning the Word Association Test, but do not indicate how information from the test should be analysed in the present study. Meaning resides in the connections between constructs (see Chapter 2). What is of interest here is whether children using concept mapping learn to generate more relevant connections to important construct terms from the domain being learned than do children leaning by more conventional means.

In the present research, the Word Association Test was compiled from a list of key construct terms from the domain of interest. (These terms were also those used in the concept maps, ensuring high validity of content.) The responses for each child were scored by recording one "link" for each occasion that one of those key terms features in a response list. Each link received the same weighting, and a symmetrical relationship would therefore register as two separate links, whereas a highly asymmetric relationship (such as in the previous example) would register only one link.

In addition to recording the number of direct links made by an individual, it is also possible to construct a graphical representation of the connections made by that person in the domain. This is achieved by laying out the constructs of interest on a sheet of paper in a suitable arrangement, as in a concept map. A linking arrow is then drawn from each stimulus term to each of the constructs on the page that was given as a response to the stimulus by the individual. The resulting diagram is similar to a concept map, except that the links are not labelled, and the hierarchical arrangement reflects the views of the researcher, not necessarily those of the respondent.

An example, taken from the research, is shown later in the chapter in Figure 6.2. It is quite easy to see from this representation which concepts have the most links to others in the domain, and also which are isolated or nearly so. It may be possible to identify distinct clusters of concepts. Hence, this form of representation is useful for exploratory purposes.

Further information on cognitive structure

It is not just the availability of concepts relevant to understanding a domain that is of interest in the present study. The "quality" of the relationships concerned is also of interest. As stated earlier, the Word Association Test does not reveal the nature of links between concepts, only whether they exist. Hence it was considered necessary to buttress interpretation of the word association data by obtaining further evidence that appropriately meaningful connections were being accessed.

The constraints on obtaining such evidence would be similar to those already discussed above in relation to the Word Association Test: that it should as far as possible be a relatively "pure" measure of cognitive structure, and depend as little as possible on general reasoning and problem solving ability. According to Gagné & White (1978), the most straightforward measure of the retention of propositional learning in a domain is free recall of sentences, which are then judged for correctness of propositional meaning. In adapting this approach for the present study, it was felt necessary to restrict the scope of sentences that respondents could recall, in order to keep the task manageable. It was also desirable to ensure that the measure prompted respondents sufficiently to ensure that, if they held a proposition, then it would be elicited.

The method used was to give the children a list of key construct terms and ask them to write sentences linking the terms. The list was the same as that used to construct the Word Association Test. However, the results, of what is still a very open task in this form, would depend to an unacceptable extent on the approach taken by respondents in selecting pairs of terms to link, as those who started with the first concept in the list and then worked systematically through the remainder trying to forge links, would perform on the task very differently to those adopting a less systematic approach. The solution adopted was to focus the task still further. First, a set of (in this case, ten) construct terms from the domain were listed at the top of the test paper. Respondents were then required to write as many sentences as they could about a fixed and limited number of constructs, using the words provided in this list. In essence, the task invites respondents to communicate any propositions they hold linking the key terms of the domain. The name adopted was Proposition Generation Task. The task used in the present research is reproduced in Appendix B.

The information required from the Proposition Generation Task was the nature of the propositional relationships between terms probed in the Word Association Test. The test scripts were therefore examined first of all to locate each instance of one of these target terms incorporated within a proposition. Then, the script was re-examined from the viewpoint of every possible combination of two terms. If a valid relationship (compatible with the textbook in use in the class) was shown for any given pair, a score of one was awarded. Trivial relationships (such as "acid and alkali both begin with the same letter") would not qualify for a mark, but in the event, none arose. The only exception allowed to the requirement that the target term be used in the proposition was when a pronoun was used which clearly referred to one of the constructs concerned.

Instruments used

A Word Association Test was produced for the topic. This contained the terms for the main constructs identified for the topic. Each term was printed ten times down the left hand side of a page in the test booklet in clear lower case lettering. Next to each occurrence of the stimulus term was a line on which to write a response, allowing a maximum of ten responses per term. The order of presentation of the stimulus terms was random, except for the first term. As the first term presented was likely to have a strong influence on how respondents constructed a context for the task, this term was chosen to be highly suggestive of the topic. As a further context setting device (after the style of Cachapuz & Maskill, 1989), pupils were given a brief problem from the domain to consider prior to taking the Word Association Test proper. The full script for the Word Association Test, including the problem, is reproduced in Appendix C. Respondents were given one minute for each term to generate responses. The same Word Association Test was used both before and after the experimental treatment.

It was not practical to retest subjects using the Word Association Tests, hence it was not possible to estimate the stability of scores from this measure. Previous work (Shavelson, 1972) suggests that word association data are relatively stable on retest when there has been no intervening instruction on the topic. Validity of content was ensured by deriving the stimulus terms directly from the text that formed the basis of the children's learning experiences.

The Proposition Generation Task written for the topic is reproduced in Appendix B. Content validity was ensured by using central constructs of the topic as the focus for the task, and inviting links to the other main constructs in the domain, which were identical to those used in the Word Association Test. Again, as no retest was possible, data on score stability are not available for the Proposition Generation Task. To avoid disruption, the Proposition Generation Task was used at posttest only, when it was expected to generate the greater amount of information. Consequently, evidence regarding the meaningfulness of the links made by the children at pretest is more limited.

Operational hypotheses

Having determined how data on cognitive structure may be obtained, it is possible to formulate specific operational and statistical hypotheses.

H_r: *The sample of Year 6 pupils who carry out concept mapping (as described above) in a science domain will show significantly higher gains in the total number of links made in a Word Association Test representing the major constructs in the domain than will the sample of Year 6 pupils who do not carry out concept mapping.*

This results in the corresponding null statistical hypothesis:

H_0: For the population, *concept mappers show equal gains to non-concept mappers in the number of links made in the Word Association Test.*

A two-tailed test of this hypothesis was required, as it could not be assumed that differences would only lie in one direction, and clearly a difference in the opposite direction to that expected would be of great practical significance. The level of statistical significance of any measured difference which would be required to reject the null hypotheses was set at $\alpha = 0.05$.

A number of further questions of a more exploratory nature were of interest. The first is that concerning the relationship between the aspects of cognitive structure probed by the Word Association Test and the Proposition Generation Task respectively. A substantial positive correlation between the two measures would suggest that they were tapping essentially the same aspect. On the other hand, a low correlation would suggest that the two aspects were dissimilar.

The next question concerned the pattern of the links made between concepts probed in the Word Association Test. It was expected that the pattern would "make sense" in terms of the actual relationships between the constructs in the scientific domain. Constructs central to the domain would be expected to show more relationships than those on the periphery. On the other hand, a lack of any meaningful structure would suggest that the measure might be deficient in some way.

Lastly, it was important to note whether the pupils in the concept mapping group were able to construct appropriate maps of the domain, showing a hierarchical structure and evidence of both progressive differentiation and integrative reconciliation over the period during which learning of the topic took place. The raw material for answering this question would be the pupils' maps constructed before and after learning.

Statistical treatment of the data

In the analysis of data from the Word Association Test, it was necessary to use some means of reducing the limitations imposed by the non-equivalent groups adopted for the research. Non-equivalent groups pose particular threats to validity in the present study. Most standard tests of statistical significance have as a prerequisite random selection of subjects, and either random assignment to groups, or assignment according to known covariate scores (Huitema, 1980). However, one analysis procedure, Standardized Change-Score Analysis (SCSA) has been developed to go some way towards overcoming the limitations of non-equivalent group designs (*ibid.*). This procedure is sufficiently respected that it became the required procedure for evaluation of large scale compensatory education programmes in the USA (Linn, 1988).

With a fully randomized pretest-posttest design, it may reasonably be expected that growth rates for scores on the test would not differ significantly between groups, unless there is a group treatment effect. With non-equivalent groups, this assumption cannot be made; indeed it is highly probable that groups *would* differ in their growth rates. It is not therefore possible to conclude that the existence of a statistically significant difference is explainable solely by differences in treatment. SCSA is founded on a "fan spread" growth model, which assumes that growth is cumulative, and therefore that growth rates for subjects with high pretest scores will be higher than growth rates for those with low pretest scores. If plotted, the growth lines would tend to spread out, fan-like. A characteristic of such a growth pattern is therefore an increase in variance on the posttest over the pretest scores. This increase leads to bias in conventional analysis of covariance when applied to the gain in score from pretest to posttest (*ibid.*). Huitema suggests a test to evaluate the fit of the fan spread model, and this is presented in Appendix D.

SCSA is an analysis of variance procedure, the object of which is to remove differential growth that is not due to treatment effects by forcing pre- and posttest variances to be equal. This is achieved by standardizing pre- and posttest scores (see Appendix D for details).

All statistical analyses except the homogeneity of variance test were carried out using SPSS/PC computer software (Norusis, 1988).

Other analyses

The Proposition Generation Task was administered at posttest only, and so analysis of gains was not appropriate. Its main purpose was to support interpretation of the word association data by way of providing concurrent validity evidence. Hence the appropriate analysis is a correlation between the two variables, which indicates the strength of the relationship between the observed values for the dependent variable (Proposition Generation Task score) and the values predicted by regression on the independent variable (Word Association Test links). As both may be regarded as interval scale variables, the Pearson product-moment correlation coefficient was used (Norusis, 1988).

In addition to the confirmatory analyses described above, various exploratory analyses were undertaken to gain insight into the functioning of the measures.

6.1.4 Implementation

On first meeting the pupils in each group, the purpose of the research was explained briefly, and their cooperation sought. The initial step with the experimental group was the teaching of the technique of concept mapping. This was carried out by the researcher, following exactly the procedure described in Appendix A. During the construction by the pupils of their second practice concept map, their performance was monitored by observing their work to ensure that they were able to follow the procedures, and that their maps exhibited an appropriate hierarchical structure. Both aspects of the children's performance proved satisfactory, evidenced by their ability to produce unaided concept maps that showed evidence of superordinate relationships.

The second session with the experimental group took place three weeks after the initial visit. On this occasion, the group took the Word Association pretest. Immediately following this, the pupils in the group were each given a strip of paper on which were printed the same construct terms as were used in the test. The children were then asked to work in their group to construct a concept map showing how they saw the terms as related. Each child made an individual copy of the map.

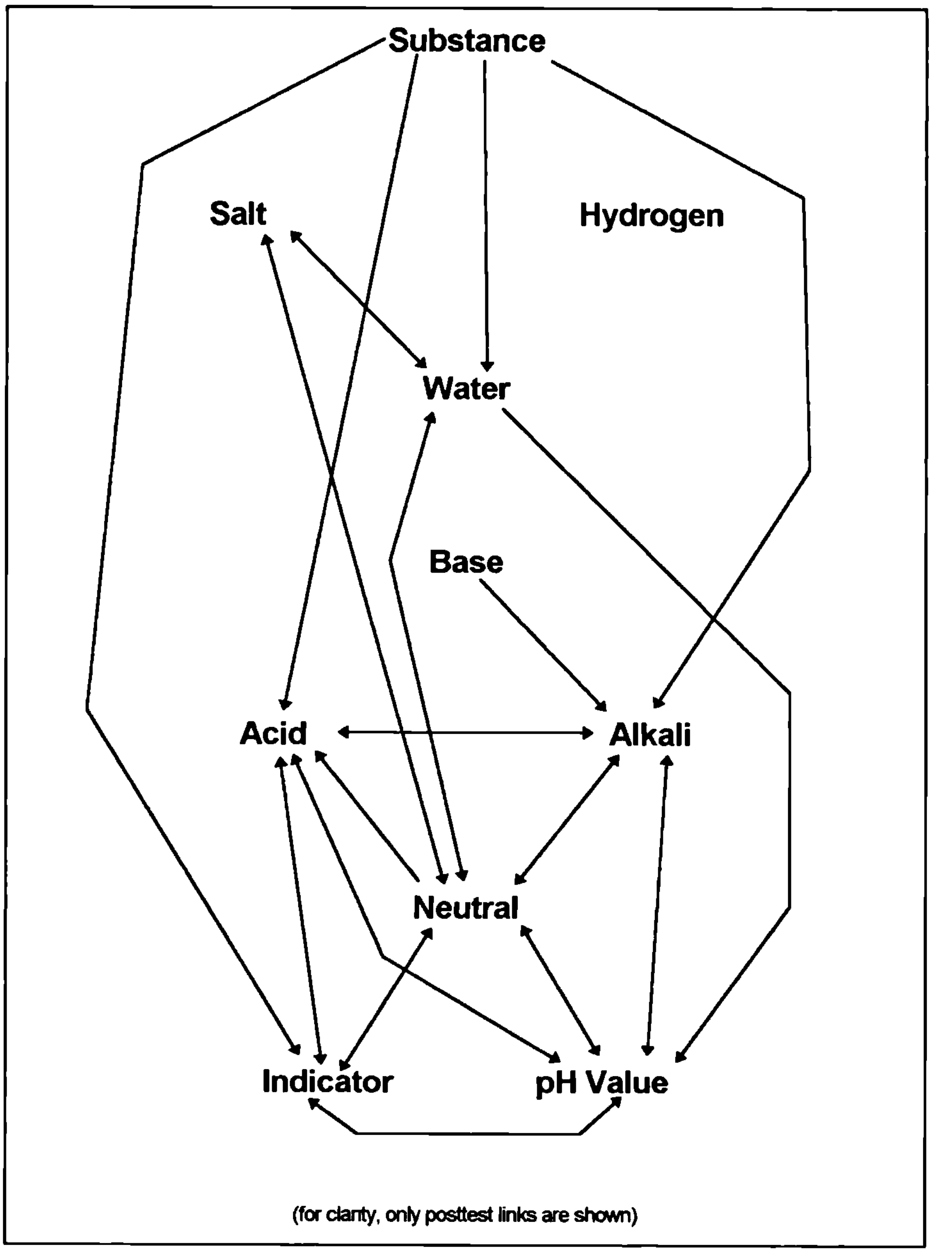
The first session with the control group consisted of administration of the Word Association Test, following exactly the same procedure as for the experimental group. Four members of the group were absent on this occasion. This was a serious loss, so an alternative date was arranged for four days later, when the pretest could be administered to the remaining children before teaching of the topic had commenced. This was accomplished without further setback.

Following administration of the Word Association pretests, teaching proceeded on the topic, according to the teacher's scheme of work. The teacher was aware of the need to maintain as much comparability as possible between the experiences of the two groups. However, it was impractical to monitor the extent to which this was achieved. At the end of the teaching unit, the pupils in the experimental group were asked to construct a final concept map, making use of the same construct words as for the initial map, but, as before, adding any further terms considered relevant.

Following this, both classes together were given the Word Association posttest. The intervening period was three days. The posttest followed the same protocol as the pretest, and the test booklets used were identical. The Proposition Generation Task could not be administered on the same occasion, and the first convenient time following the Word Association posttest was used. The intervening period was 14 weeks. Effectively, this means that the Proposition Generation Task served as a delayed posttest, and therefore was used to provide predictive rather than concurrent validity evidence.

6.2 Results

Figure 6.2: A Map of Associations made in the Word Association Test



The first source of data was the Word Association pre- and posttests. These yielded a score, and in addition a schematic "map" of the cross-links between the ten concepts was prepared for each pupil. This enabled an overall impression of the pattern of links to be gained. An example of such a map is shown in Figure 6.2.

The second source of data was the Proposition Generation Task, which was administered at posttest only. An example is shown in Figure 6.3.

Figure 6.3: A Completed Proposition Generation Task

The Acid Test

Use these terms in your answers:

---

acid alkali base hydrogen indicator neutral pH value salt substance water

---

1. Write as many sentences as you can about **acids**, using the words above
2. Write as many sentences as you can about **alkalis**, using the words above
3. Write as many sentences as you can, using the words above, to describe **what happens when an acid and an alkali are mixed**

1. Acid has a PH value of bellow 7. Acid is a class of substance. An example of an acid is hydrogen. Acid comes at redish on indicator.

2. Alkali has a P.H value of about 8 to 12. Alkaline is a class of substance. An example of an alkaline is salt. Alkaline goes purple in indicator.

3. when equal amounts of acid and alkali are mixed they turn nuetral. Nuetral has a P.H value of 7. Nuetral is a class of substance. An example of neutral is water. neutral goes green in indicator.

I can't remember anything about base

(Redrawn for clarity)

The third source of data was the concept maps constructed by the pupils in the experimental treatment group. Examples are presented further below in Figures 6.8 and 6.9.

6.2.1 Effects of Concept Mapping

The first variable to be examined statistically was the number of links made by respondents between the terms in the test. The medians, interquartile ranges and ranges of the data for both pre- and posttest are depicted in Figure 6.4. This figure illustrates the changes in the number of links made by both groups over the period of the experiment, and shows a steeper rise for the treatment group than for the control group. The means and standard deviations for the same data are reported in Table 6.1, together with a 95% confidence interval for the mean, calculated from the standard error.

Figure 6.4: Schematic Plot of Pre- and Post-topic Word Association Test Scores

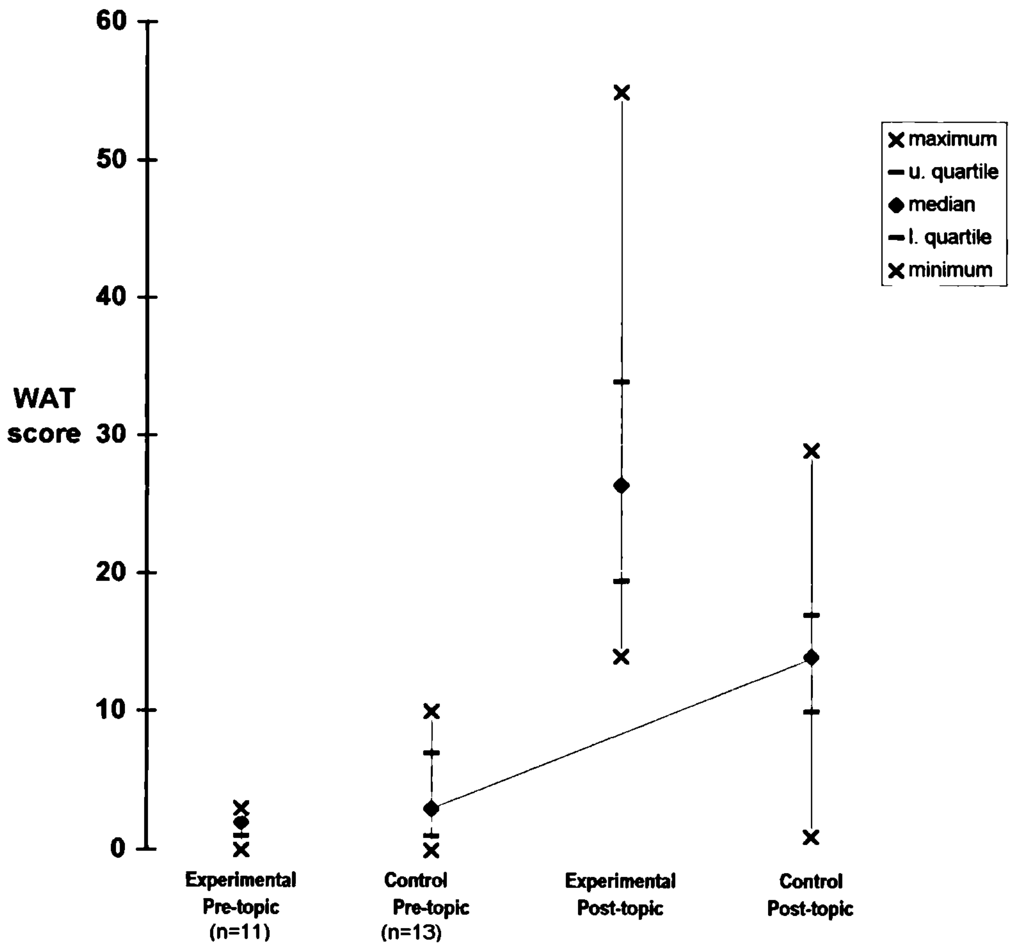


Table 6.1: Total Number of Links (Word Association Test)

Group	Pretest Mean	St. Dev	95% Conf. Int.	Posttest Mean	St. Dev	95% Conf. Int.
Experimental (n=11)	1.8	1.1	±0.6	28.7	12.3	±7.6
Control (n=13)	3.9	3.4	±1.8	13.8	6.9	±4.1

The low distribution of scores meant that there was the danger of a "floor" effect on the pretest. This in turn would imply that the information obtained in the pretest might not discriminate adequately between individuals, and would therefore be of limited value in removing any effects due to non-equivalence of the two groups.

At pretest, a between-groups comparison was made using a distribution-free statistical test; the Mann-Whitney U test (Norusis, 1988). There was no statistically significant difference between the distributions of scores for the two groups[10]. Thus the hypothesis that the two groups were equivalent on this variable at the beginning of the experimental run was not rejected.

The median number of links made (two for experimental group pupils and three for control group pupils) was very low. This suggested that the children knew little about the domain initially, and inspecting the content of the response lists supported this view. Only two children, for example, (both in the control group) appeared to have attached any meaning to the term "pH value". Other terms within the domain tended to be associated with experiences outside the domain. Thus, both "acid" and "alkali" prompted the response "battery". "Base" was associated with the idea of "headquarters", and "indicator" with cars. "Salt" and "water" were often connected, presumably because of experience of salt water in the sea: there was nothing in any of the response distributions which suggested the scientific meaning of "salt". However, there were some indications that the control group were more knowledgeable about the domain at pretest than the experimental group. For example, one respondent in the control group gave "not acid" and "not alkali" in response to "neutral". No pupil in the

[10] $p = 0.31$

experimental group gave any response in the pretest which suggested a scientifically relevant link between two of the stimulus terms.

At posttest, many more links were generated, and there was a large difference between the two groups in the number of links. In order to establish whether the learning gains were statistically significant, the non-parametric Sign test was used (Norusis 1988). This enabled the rejection, for both groups, of the null hypothesis that the posttest scores were equal to the pretest scores[11]. Inspection of the response distributions suggested that the children had a much wider range of scientific meanings for the stimulus terms at posttest.

The distributions of responses at posttest were similar to normal, and so comparison between the groups by parametric statistical tests was possible. In order to determine whether the data conformed to the fan spread model, the homogeneity of variance test described in Appendix D was applied to the pre- and posttest scores. The hypothesis that the pretest and posttest variances were equal was rejected[12]. Use of standardized change-score analysis to compare the gains for the two groups was therefore appropriate. The mean adjusted gain score for the experimental group was 20.31 (standard deviation = 12.21) and for the control group it was -1.05 (standard deviation = 8.60).

The results of an analysis of variance on the adjusted gain scores are reported in Table 6.2.

Table 6.2: ANOVA on Adjusted Gain Scores

Source of Variation	Sum of Squares	DF	Mean Squares	F
Main effects (Group)	2390.571	1	2390.571	20.680
Explained	2390.571	1	2390.571	20.680
Residual	2080.736	18	115.596	
Total	4471.306	19	235.332	

Critical F (0.01, 1, 18) = 8 28

[11] $p < 0.002$ for both groups

[12] $t = 9.06$ for 17 degrees of freedom. The critical value for $\alpha = 0.30$ is 1.067

The resulting *F* for the main effect of concept mapping treatment was greater than the critical value, and therefore statistically significant. This demanded rejection of the null hypothesis that the two group mean gain scores were equal. Thus, an effect attributable to the concept mapping was indicated, and the difference in mean gain shows this to have been substantial. However, this result needs to be treated with the utmost caution, as there are rival explanatory hypotheses that cannot be discounted. Firstly, the possibility of a floor effect on the pretest data meant that pre-existing group differences may not have been adequately controlled. Secondly, there was no way to be certain that the two groups were treated entirely equally. Also, the possibility initially planned of crossing over the two groups for a second experimental run was precluded.

Proposition Generation Task

Figure 6.5: Schematic Plot of Proposition Generation Task Scores

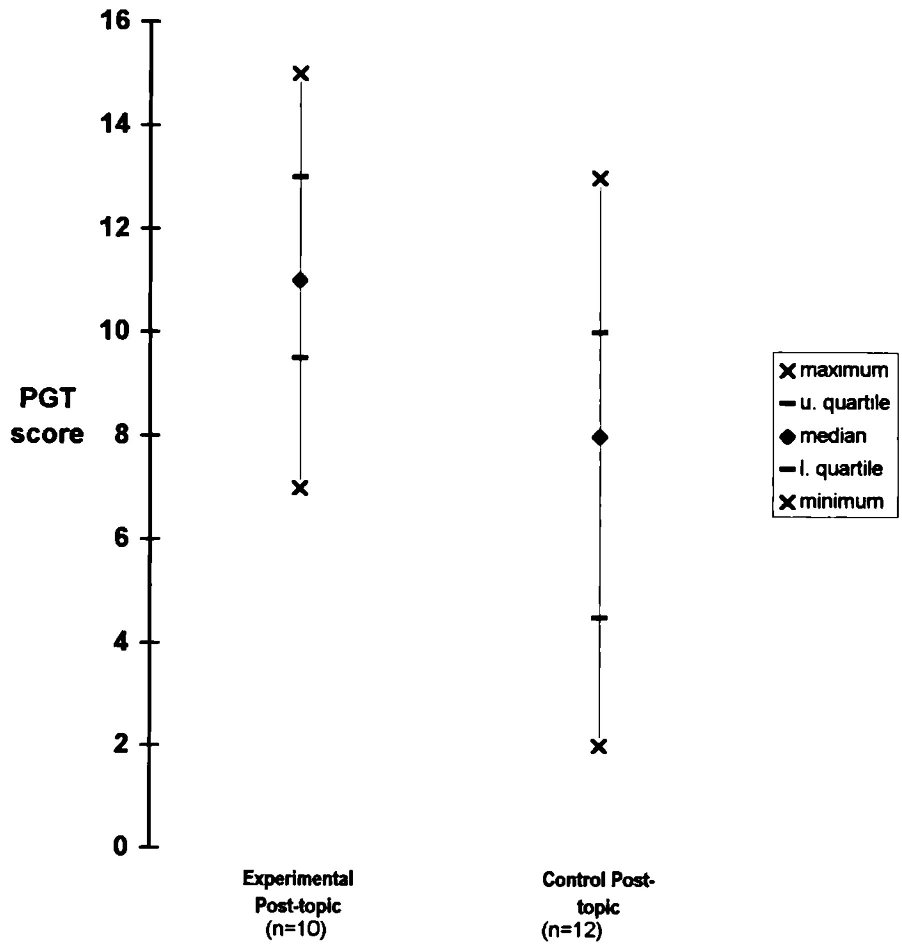


Figure 6.5 shows the median, interquartile range and range of the number of valid propositions generated in the Proposition Generation Task. As with the number of links, the mean number of propositions generated differed between the two groups. Table 6.3 shows the mean, standard deviation and 95% confidence interval of the scores for the two groups.

Table 6.3: Total Number of Valid Propositions (Proposition Generation Task)

Group	Mean	St. Dev.	95% Conf. Int.
Experimental (n=10)	11.1	2.5	±1.6
Control (n=12)	7.6	3.3	±1.8

The confidence intervals here suggest that the two mean scores were significantly different from one another. But as there was no means of taking account of possible pre-existing group differences with these data, it cannot be concluded on this basis that the difference was attributable to the effect of concept mapping. However, this was not the main purpose for administering the Proposition Generation Task. Rather, it was to aid interpretation of the Word Association Test data.

There was a substantial correlation between the number of relationships generated in the Proposition Generation Task and the number of links in the Word Association Test posttest[13]. In order to examine this relationship more closely, the propositions variable was plotted against the links variable. This plot, shown in Figure 6.6, indicates that in addition to the large spread of data there appears to be a curvilinear relationship between the variables, with the number of propositions tailing off at higher values of the links variable. This is suggestive of a "ceiling" effect in the Proposition Generation Task. It was decided, on the basis of this curvature, to try transforming the variables in order to obtain a linear plot. Taking the square root of the Word Association Test score led to a better linear fit, but only a small increase in the correlation between the two sets of scores[14].

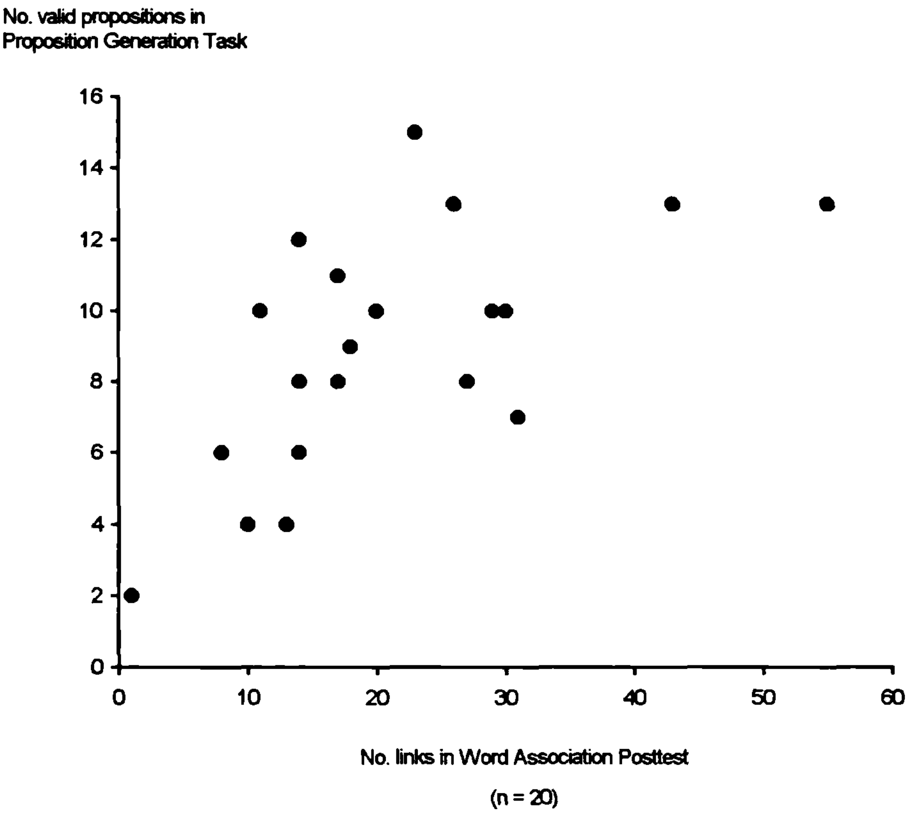
Such a ceiling effect is unsurprising, as the Proposition Generation Task is more demanding in terms of cognitive processing than the Word

[13] $r = 0.63, p = 0.003$

[14] $r = 0.69, p = 0.001$

Association Test, and is more likely to be curtailed through shortage of time. It is also possible that children thought some links too obvious or trivial to be worth expressing in the Proposition Generation Task which nevertheless they registered in the Word Association Test: that acids and alkalis are substances, for example. The Word Association Test could well (and often did) register a relationship between concept pairs that were only related indirectly, via another concept, in the Proposition Generation Task.

Figure 6.6: Plot of Proposition Generation Task Score with Word Association Posttest Score



These data show that scores on the Word Association Test were predictive of scores on the Proposition Generation Task administered several weeks later, and consequently support use of the Word Association Test as a measure of a significant aspect of cognitive structure. However, the relationship is not especially strong. This may in part be due to the different characteristics of the measures used and to the time interval. But the differences may also be due to the two measures' to some extent tapping distinctive aspects of the construct being addressed.

Patterns of Associations between Concepts

Schematic "maps" were prepared of the associations made by pupils in response to the Word Association Test. An example of such a map is presented above, in Figure 6.2. Examining these maps gave an overall impression of the patterns of links made by the pupils. The first such impression is that certain terms regularly attracted more connections than the others. These were "acid" and "alkali", followed by "pH value" and "neutral". This makes good sense, in view of the title of the topic, and its main teaching focus. Pupils in the concept mapping group tended to make more links to these key constructs than did those in the control group, suggesting that the greater total number of links they made was not simply random "noise", but a reflection of the relationships within the topic, as they experienced it.

Figure 6.7 shows a detailed comparison of the frequencies with which the two groups linked each pair of terms in the test. In the figure, each cell represents a conjunction between two of the stimulus terms used in the Word Association Test. The upper tally shows the frequency of links made by the experimental group, and the lower tally the links made by the control group.

This analysis showed that the differences between the two groups in the frequencies of links made between pairs of terms was not evenly distributed among the cells. Some pairs were linked with similar frequency by both groups, while others showed much greater differences. "Water" was the term exhibiting the widest differences overall. Examining the response lists produced by the children showed that, while the children in the two groups made similar numbers of responses overall for the term, those in the experimental group made rather more within-domain links (for example, that water is neutral, that it has a pH of about 7) than did the control group. The figure also reveals that the experimental group more often made links interrelating the terms "base", "acid" "alkali" and "substance", suggesting greater integration of these ideas. Mostly, the control group focused on the connection between "acid" and "alkali", while very few showed any sign of a scientific meaning for the term "base". The teacher had introduced this term to the children, but had not stressed its significance, and it seems that it was not generally integrated into the topic by the pupils. This was reflected in the Proposition Generation Task, where very few children incorporated the term in a valid proposition. Indeed, a misconception seems to have arisen that a base is a neutral substance, and some concept maps constructed by pupils in the experimental group also revealed this relationship.

Figure 6.7: Comparison of Post-Topic Links between Constructs for Experimental and Control Groups

	ACID	SALT	ALKALI	SUBST	BASE	pH	WATER	INDIC	HYDRO	NEUTR
NEUTR	IIIIII IIIII	IIIIIIII IIII	IIIIIII IIII	I III	III I	IIIIII IIIIII	IIIIIIII IIII	III III	II	
HYDRO	IIIII III	I	III	II II	I		I I			
INDIC	IIIIII IIIIII	I	IIIIII IIIII	IIII I	II	IIIIIIIII IIIIIII	I I			
WATER	II III	IIIIIIII III	II I	IIIIII III	II I	IIIIII III				
pH	IIIIIIIIII IIIIIIII	II II	IIIIIIIII IIIIIII	 II	II I					
BASE	IIIIIII II	II	IIIIIII I	III I						
SUBST	IIIIIIII IIII	IIII II	IIIIIIII IIII							
ALKALI	IIIIIIIII IIIIII	III I								
SALT	II II									
ACID										

Frequency of Links (corrected*)

Experimental Group (Top)
V.
Control Group (Bottom)

* Since the numbers of children differed between the groups, the frequencies shown are corrected for a group of ten children

Both groups provided a range of responses to the term "neutral" that suggested a grasp of its scientific meaning, but rather more of the experimental group included the substances "water" and "salt" in their lists. "Hydrogen", on the other hand, was another construct that was rarely integrated with others. Again, this isolation was reflected in the Proposition Generation Task responses, where the relevance of hydrogen to the chemical structure of acids and water did not feature in any response.

6.2.2 The Pupils' Concept Maps

Concept maps were constructed by pupils in the treatment group, both at the beginning and at the end of their encounter with the topic "Acids and Alkalis". Copies of the maps were made and retained, and these provided evidence regarding whether the pupils engaged in the task of concept mapping in the manner intended.

Hierarchy

Almost without exception, the maps produced showed a meaningful conceptual hierarchy. In all of the pre-topic maps, "substance" was placed at the top, and there were often identifiable strata of differing generality. Thus "salt", "water" and "hydrogen" were typically placed at the same level, and "acid" and "alkali" were sometimes together at a different level. Post-topic maps tended to show a clearer hierarchical arrangement, normally subsumed under the construct "substance". However, one group chose to put "base" at the top, with validly labelled links to "acid" and "alkali" below. This group's maps did not show any obvious hierarchy, although it is perfectly possible that the pupils saw this as a reasonable hierarchical arrangement.

Progressive Differentiation

Two maps, by the same pupil, illustrate this section. Figure 6.8 shows the arrangement of constructs in the pre-topic map. Figure 6.9 shows the arrangement of the same constructs after teaching on the topic.

Figure 6.8: A Pre-Topic Concept Map

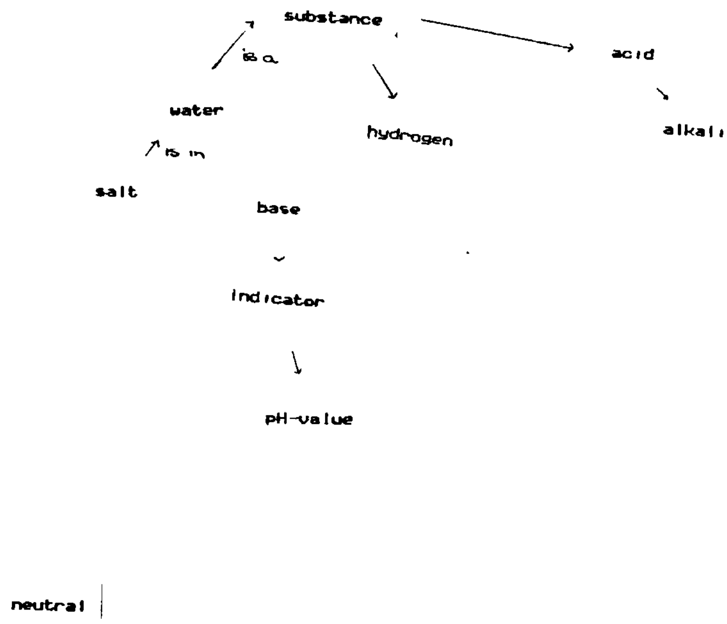
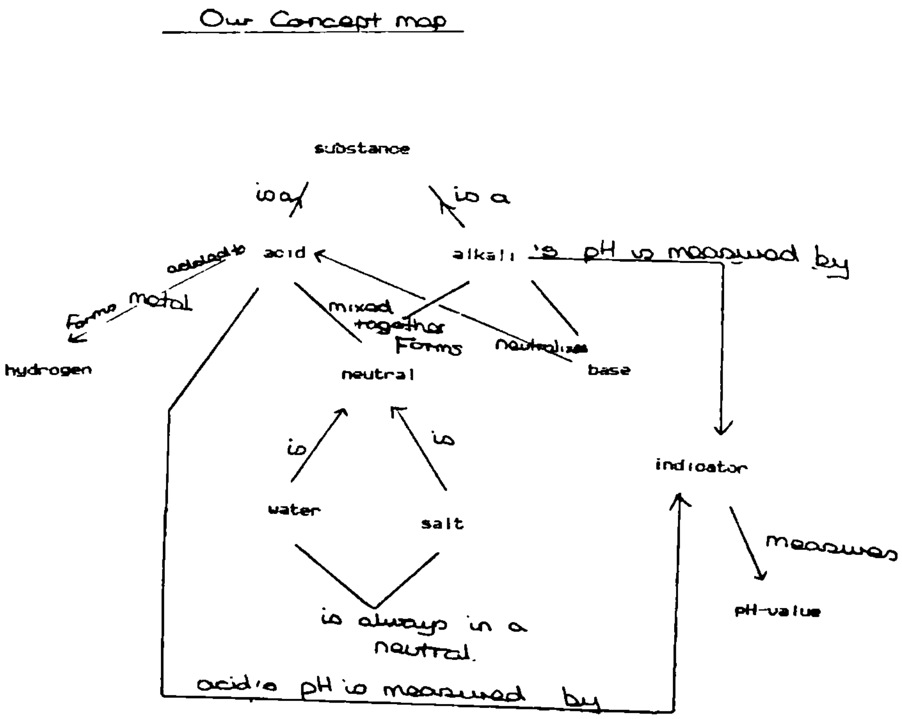


Figure 6.9: A Post-Topic Concept Map



The first map shows only two connections with "acid", neither of which is labelled. The arrangement is suggestive of some understanding of acids as substances. However, the nature of the link to "alkali" is not known. A number of other pupils at this stage believed (or guessed) that an alkali is a type of acid, and this may be the case here too. Thus, the scientific meaning attributable to both these terms at the outset appears to have been minimal. By contrast, the later map shows a number of valid propositions incorporating the terms "acid" and "alkali". There has clearly been progressive differentiation of the two concepts over the period of the study. The pupils were encouraged to add further terms of their own to the maps, and in this case "metal" has been linked appropriately into the structure.

Probably, this was done in order to integrate the construct "hydrogen" into the map. This pupil appears not to know the significance of hydrogen to the concept of "acid"; however, the structure created would allow further differentiation of this relationship in the future. Such a map, if retained, could prove to be a useful means of reviewing prior learning and of tying in newly introduced constructs.

Integrative Reconciliation

The same pair of maps serves to illustrate this section too. In addition to the cluster of terms linked with the superordinate construct "substance", there is also in the earlier map a small chain of constructs which includes an unlabelled link between "indicator" and "pH value". This may be guesswork, or it may be a vaguely remembered connection; the Word Association pretest for this pupil did not suggest any scientific meanings for these terms. In the post-topic map, not only is there evidence of a better established link between these two constructs, but also they are no longer separated from the concepts of "acid" and "alkali".

The second map also shows a more direct relationship between the terms "hydrogen" and "acid". In the earlier map, this was via the superordinate construct "substance" only. In this later version, another substance ("metal") brings a different structural relationship between "hydrogen" and "acid". Presumably, although not explicitly shown, the pupil in question has not ceased to think of hydrogen as a substance.

This pair of concept maps, drawn before and after teaching, reveals therefore several recombinations among existing concepts; what Ausubel (*et al.*, 1978) termed integrative reconciliation. Together with the evidence for progressive differentiation and a more ordered hierarchy of concepts, this shows considerable conceptual development. Such development was in evidence with other pupils as well. In fact, as this example suggests, it may be that development was such that it sometimes became difficult to consider and incorporate all the possible relationships known to exist between the ten constructs in the map. Several relationships (with the concept "substance" for example) have probably remained tacit. This pupil also resorted to using the same terms ("pH", "neutral") in more than one position on the map, suggesting that finding an appropriate arrangement proved challenging.

The question remains as to how performance in the collaborative concept mapping task related to the children's performance in the Word Association measure. Because there was little variation in the number of links made in the concept maps, and because the number of children who made concept maps was small, it was not practicable to carry out a correlation analysis. Instead, an exploratory comparison was made between the links made in the concept maps and the patterns of connections identified in the Word Association Test responses. This was carried out for the post-topic session only, as there were very few links made in the pre-topic test. Figure 6.10 shows the result of this analysis. In the figure, each cell represents a conjunction between two of the terms that were common to both concept map and Word Association Test. The upper tally shows the number of cases for which both measures agreed in registering a linkage (or non-linkage) between the pair of terms. The lower tally shows the number of cases for which the one measure registered a link but not the other.

The tallies show that the two sets of data agreed in the large majority of cases. However, there were a few cells for which there was substantial disagreement (considered to be when the number of disagreements equalled or exceeded the number of agreements). For these cells, mostly the discrepancy was where links were made in the Word Association Test that did not correspond with a link in the concept map. These discrepant cases were examined further, in order to discern the source of the disagreement and to identify whether there might be a common factor involved.

Figure 6.10: Relationship between Concept Map Outcome and Word Association Test Outcome

	ACID	SALT	ALKALI	SUBST	BASE	pH	WATER	INDIC	HYDRO	NEUTR
NEUTR	IIIIIII II	IIIIIIIII	IIIIII III	IIIIIIII I	IIIIII III	II IIIIIII	IIIIIIIII	IIIIII III	IIIIII III	
HYDRO	IIIII IIII	IIIIIIII I	IIIIII III	IIIII IIII	IIIIIIII I	IIIIIIIII	IIIIIIII I	IIIIIIIII		
INDIC	IIIIII III	IIIIIIII I	IIIIII III	IIIII IIII	IIIIIII II	IIIIIIIII	IIIIIIII I			
WATER	IIIIIIII I	III IIIIII	IIIIIII II	IIII IIIII	IIIIIII II	III IIIIII				
pH	IIII IIIII	IIIIIIIII	III IIIIII	IIIIIIIII	IIIIIII II					
BASE	IIIIIII II	IIIIII III	IIIIIIIII	IIIIII III						
SUBST	IIIIII III	IIIII IIII	IIIII IIII							
ALKALI	IIIII IIII	IIIIII III								
SALT	IIIIII III									
ACID										

Frequency of Agreement in Outcomes

Agreement Post-WAT/Concept map (Top)
V.
Disagreement Post-WAT/Concept map (Bottom)

This analysis yielded some interesting findings. A very substantial proportion of the discrepant cases involved a small cluster of constructs: "substance", "water" and "pH value". Quite often, children had associated "substance" with specific instances such as hydrogen, water and salt in the test, but did not make a direct connection in the concept map. However the links that were made in the concept maps suggested that the children could have linked these terms together had they chosen to. They appeared to understand quite well that these were substances. In a similar way, "pH value" was often not connected in the concept map to constructs such as "neutral", "acid" and "alkali", although the network of relationships made included indirect connections that showed the children understood how these ideas are related. "Salt" and "water" were frequently associated in the Word Association Test, whilst the children did not choose to link the ideas together in the concept map. Very often, the link between these two terms was evident in the pre-topic test as well, so this is a case of a connection that was in place before the topic began, and was not due to studying the topic. This was also the case for the link between "substance" and "hydrogen", and to a lesser extent between "substance" and "water". With these few exceptions, it seems that the making of a link in the concept map predicted the making of an equivalent connection in the Word Association Test.

6.2.3 Conclusions

The results presented in this chapter enable some provisional conclusions to be drawn about concept mapping. Firstly, the children seemed capable of learning to produce appropriately structured concept maps, and this did not prove troublesome. Moreover, there is evidence that the way the children structured these maps was sensitive to the knowledge they had acquired. There was seen to be a relationship between links made in the concept map, and the connections identified in the Word Association Test, and this in turn produced a structure that was interpretable in terms of the domain being studied.

There is some evidence for a learning effect due to the concept mapping activity, as evidenced by the Word Association Test outcomes and supported by the Proposition Generation Task. This, it must be reiterated, is most tentative, for reasons that have been discussed. However, there is sufficient here to warrant looking for possible group influences on learning; the subject of phase three. Hence the focus now moves from *whether* concept mapping has an effect to *how* it does.

In terms of the instruments used, significant relationships were identified between the three sources of data, but also some differences. This suggests that the concept map, the Proposition Generation Task and the Word Association Test were addressing different, though closely related, dimensions of cognitive structure, and therefore could potentially be a useful conjunction of instruments in research where detailed information on cognitive structure is required.

Whilst there have been worthwhile outcomes from this phase of the research, there were also some practical difficulties encountered, and in consequence part of the research could not proceed as planned at all. However, there is already a large body of related work with which the preceding outcomes tend to concur.

7

RESEARCH DESIGN: PHASE THREE

7.1 Research Design for the Third Phase

This research phase consisted of two nested elements. The main study was set up to identify the processes at work in collaborative concept mapping groups, and the substudy to confirm whether collaboration had an effect on the concept maps produced. Over the course of this chapter, the design of these interrelated elements will be discussed.

7.1.1 The Substudy

The aim of the substudy was a subset of the aims for phase 3, and was guided by the research question discussed in Section 5.1.

Q3: *Does collaborative or individual concept mapping better promote the development of scientific meanings?*

The research was to be carried out alongside the main investigation (to be described next), to which it was considered complementary. To address the question, a comparison was made between the concept maps produced by individual children and those produced by children working collaboratively. This requirement imposed a degree of artificiality on the research design, as children within the same class needed to be assigned to each of the two conditions, and also, for organisational reasons, placed greater constraints on the times available for conducting the research. Because of these impositions and the attendant disruption to classroom routine, it was decided to carry out the substudy in only one of the three classes involved.

The notion "development of scientific meanings" needed to be operationalized to enable quantification. Although such quantification inevitably reduces the richness of the data, and so detracts from determining their meaning, this was considered justifiable for this element in the research, as the main part of the study was intended to explore the nature of collaborative concept mapping in greater depth.

The primary assumption on which the substudy was based was that the finished concept map was an inscription that represented the outcomes of thought processes. This led to the operational assumption that the number and nature of the relationships shown in a concept map was a direct reflection of the quality of thinking that went into producing the map, either by an individual or by a group. If collaboration really does contribute to the discussion and negotiation of scientific meanings, then we would expect the concept maps produced under these conditions, *ceteris paribus*, to show more scientifically appropriate links and fewer links based on misconceptions. Hence we can specify as a research hypothesis:

H_r: *Collaborative concept mappers will incorporate more scientifically appropriate relationships in their concept maps than individual concept mappers.*

This hypothesis was investigated for both pre- and post-topic maps. The method of deriving an index of scientific appropriateness of the links in the children's concept maps is described in 8.2.

To investigate the above hypothesis, an experimental, rather than quasi-experimental, design was developed. The requirement that statistically generalizable results should be obtained entailed random assignment. Hence classroom groupings already constituted by the teacher could not be used. The teacher and the overall teaching programme were common to all the children in the class involved, providing a good degree of control over the learning opportunities for the children in the two conditions. However, within primary classes, an element of individualized provision for learning is to be expected, and in addition to this, individual children may receive slightly different learning experiences due to such occurrences as attending special music tuition whilst others in their class are doing work in science. By randomizing assignment, threats to internal validity due to these differences (that is to "history") can be reduced, though never eliminated entirely (Cohen & Manion, 1980).

Another threat to validity internal to an experimental design is the possibility of uncontrolled factors' intervening between the experimental "treatment" and the measurement of the dependent variable. In this study, there was no time lag involved, as the dependent variable was not derived from a separate posttest, but directly from the concept maps used as part of the "treatment". This threat was therefore eliminated. For similar reasons, there was no possibility of pretest sensitization to treatment for either of the groups.

In order to meet the assumption of independence of scores underlying confirmatory statistical tests (Wiersma, 1969), the children in the collaborative groups made their own concept maps. Thus, although the discussion was collaborative, the maps were produced and then scored independently.

7.1.2 The Main Study

The research questions for the main study in this phase were as discussed in 5.1.

Q4: *What processes characterize the discussion and production of a concept map by these groups of children?*

Q5: *How do these processes relate to the broader views about learning in science developed in the first three chapters?*

Q6: *Does the emerging concept map help to structure the children's activity in a way that encourages the critical sharing of meanings and the emergence of new understandings?*

Q7: *Is the group production of a concept map best characterized as a constructive or a reconstructive activity?*

There was no developed theory to describe or explain the processes that characterize collaborative concept mapping, and so there could be no significant prior hypotheses that could be tested. The variables involved could not be specified in advance (and consequently could not be manipulated experimentally). The focus therefore needed first to be on identifying those variables in order to facilitate description, leading to the development of theory.

The research was founded on assumptions that the processes by which meanings are created and negotiated can be located in communicative practices situated within particular activities. This is the view of social constructionism, as found in the work of Wittgenstein (1967) and elaborated in Chapter 2. Further, these communicative practices are realized by means of semiotic (or sign) systems, of which language is one (Halliday, 1978; Lemke, 1990). For this phase, it was therefore important to gather data on the semiotic practices associated with settings in which the use of concept maps was related to genuine classroom purposes. The research approach most appropriate for this purpose is best described as a form of case study. The "cases" concerned were narrower than how this term is typically used, referring to *instances of the use of collaborative concept mapping*. These cases were selected from naturally occurring instances, that is, primary school classes in which concept mapping was to form an integral part of a planned learning sequence. The composition of the classroom groups in which concept mapping was carried out was determined by the teacher, in accordance with normal classroom practice, ensuring that the research reflected as closely as possible actual classroom usage. Only in the substudy did a degree of artificiality enter the classroom organization, as explained above. The rôle of the researcher was primarily that of non-participant observer, gathering data without participating in the concept mapping itself. However, as with the classroom teachers involved, the researcher interacted with the children as they engaged in the task, taking on the rôle of assistant teacher on these occasions. In each case, the children knew that they were helping the researcher to "do research". Data pertinent to the study of these cases were any instances of semiotic practice associated with the concept mapping, and any background information necessary to interpret these primary data.

Two main kinds of semiotic practice are associated with collaborative concept mapping. The first of these is the production of a concept map, which is intended to be a meaningful representation of the way pupils see a set of constructs as being related. The semiotic function of the map is not delayed until it is complete, but operates also while it is being constructed, when the given terms can be referred to by those making the map (Roth &

Roychoudhury, 1992). The second practice of interest is the talk that takes place within the group as they construct the map. It was therefore necessary to obtain data on both these forms of communication, and so as not to obscure important detail, it was necessary to obtain the data in as full a form as possible. Concept maps were therefore collected, and the talk (which included reference to the emerging map) recorded verbatim. Maps were constructed both before the teaching unit had commenced (pre-topic map), and at the end of the topic (post-topic map). There were small differences in the way that these sessions were organised from class to class, and so further details regarding each case are given below, in 7.3.

Apropos of the methodology, the hermeneutic perspective provides a guide to considering how to address the data to answer the research questions. The goal of the research, in essence, is to develop a text (a "book of science") to describe and explain the data, which in themselves require interpretation. We could develop any number of ways to describe the data, but the data will not tolerate just any description. The constructs in the description must *refer* in some consistent way. The starting point for the description must be preconceptions. These could be preconceptions about potential categories imposed from outside the discourse (as is the case with systematic observation systems). Alternatively, they could be about the meaning of whole interactions encountered within the discourse, grasped intuitively as from the point of view of a native speaker of the language (an approach that has much in common with ethnography; Goetz & LeCompte, 1984). However, what the hermeneutic perspective suggests is that, whilst either of these might be a starting point, neither is the end point. There must be an iteration between part and whole, in order to enter the discourse and apprehend its meaning more fully. A fruitful way forward would therefore be to employ *both* perspectives, in the expectation that each one may be enriched by viewing it through the other. Analysis of the data was therefore carried out post-hoc, and cycled between prior theoretical schemes and intuitions about the meaning of episodes in order to arrive at a category system. The development of the category scheme and the assumptions underlying this development are described in detail in Chapter 8.

Accepting the reservations expressed above about the rôle of prior conceptions, Strauss' & Corbin's (1990) grounded theory method provided a framework by which the analysis might proceed. In this method, the analysis moves from the emergent category scheme (the "open" coding), through the establishment of causal/temporal relationships ("axial coding"), to the development of a validated "story line" to account for the processes at work in the groups. This was of necessity a qualitative analysis in the first instance. Quantitative analysis, as one means of summarizing patterns in the data, was used, firstly where it could help to identify potential relationships ("exploratory" data analysis) and also where it could provide "confirmatory" evidence to validate the relationships identified (see Erickson & Nosanachuk, 1979).

The children's actual concept maps were collected in some cases, and accurate copies made when this was not possible. Children's discussion was audio-taped, as this provided an exact record of what was said. This method was chosen in preference to video-taping in order both to ease resourcing problems (it would not be practical to have more than one group video-taped in a session) and to minimize obtrusiveness. The price paid for this choice was the loss of data on non-verbal communication in the group which might, in a very small number of instances, have rendered interpretable what was otherwise difficult or impossible to classify. Often, background information collected reduced this disadvantage. This background information consisted of the teachers' plans for the scientific work covered in the topic; the sets of terms given to the children to use in their maps; details of the groupings of children within the class; and field notes of unstructured observations and any additional information that would clarify events recorded in the primary data.

7.2 Sample

The population of interest in the present research was children in upper primary school classes, which are defined for the purposes of this research as Years 5 to 6 in English schools (9- to 11-year-olds). The nature of the sample selected is of interest in establishing to which population the findings of the research may be generalized. The need to study concept mapping in naturalistic settings precluded drawing samples at random from the entire population with the attendant need to instruct the children how to make concept maps and to train teachers in incorporating concept mapping into their work in science. Consequently, teachers were contacted who were known to be using concept mapping in the course of their teaching.

This resulted in selection of the following:

- one class of 32 mixed Year 5 and 6 children in a suburban junior school, situated in a moderate to high income area. The children were almost entirely from the ethnic majority, with English-speaking backgrounds. The children were taught by the deputy headteacher, who had no specialist training in science;
- one class of 24 Year 6 children in an inner-city primary school, situated in an area of low prosperity. Approximately one quarter of the children were from minority ethnic groups, most of whom learned English as an additional language. The teacher had five years' teaching experience, and had received specialist training for the teaching of primary science;
- one class of 28 Year 5 children, in the same inner-city school, and with a similar mixture of cultural backgrounds. Their teacher was the teacher who taught the above class.

In each case, a wide spread of academic abilities was represented within the class. Although not selected at random, there was no reason to suspect that these were anything other than typical classes for the areas described. The contrasting backgrounds of the children helped to broaden the range of settings to which the results could be generalized. In all three classes, concept mapping had been used primarily to assess understanding of science.

The first of the above classes was used for the substudy. For the purpose of assigning them to the two experimental conditions, the children's teacher first ranked the children in order of attainment in science (based on her ongoing records, which she was not willing to disclose). Working down the ordered list, the first pair of children were assigned at random, one to each of the two conditions, then the next pair were assigned in the same way, and so on. This had the effect of creating a stratified random selection within the class. The children in the collaborative condition were then put into working groups by their teacher.

The resulting subsamples were:

- Individual condition: 8 boys ; 7 girls
- Collaborative condition: 9 boys ; 6 girls (groups 1-4)

Two additional children were absent from class, and were not included in the assignment. The collaborative groups consisted of three mixed-sex groups of four children and one of three children.

For each of the other two classes, the children worked collaboratively in their usual classroom groups, as assigned by their teacher. Five groups were selected in each class to be the focus of the research, in accordance with the recording equipment available. In each case, this left one group excluded, selected at random. Groups 5 to 9 (19 children altogether) were in the second class, and groups 10 to 14 (23 children) in the third.

7.3 Implementation

A distinctively different science topic was studied and mapped by each of the three classes, referred to hereafter as topics 1 to 3. The topics were chosen by the teacher in each case, and no attempt is made here to justify either selection of the topics or their content.

7.3.1 Topic 1

The focus was "Habitats", drawn from the then extant Programme of Study for Life and living processes (Great Britain, DES, 1991). The children observed two contrasting habitats, and learned about: how organisms are suited to their environment; how they compete for resources; and food chains and their relationship to habitat. The terms to be mapped were in this case chosen by the researcher, in conjunction with the teacher. They were:

habitat; energy; competition; predator; food chain;

plant; animal; food; sunlight; survival.

On each occasion, the children were given these terms on a strip of paper, which they could tear up and move around before fixing the position of the terms by writing them on their sheet of paper. They were told that they could add terms of their own if they wanted to. A first concept map was produced by all the members of the class at the same time, prior to any work in relation to the topic. One map was constructed by each child in the class, and these were all collected for later analysis. Tape recordings of the discussion were attempted for the four collaborative groups by means of a single compact tape recorder placed centrally for each group. However, two tapes failed, one due to the tape's becoming entangled in the recorder mechanism, and the other due to a faulty microphone connection.

Four-and-a-half weeks later, the post-topic concept mapping session took place, following the end of work on the topic. Two children from the individual condition were absent for this session, and so their data were removed entirely from subsequent analyses. Again, each child completed one concept map, and these were collected for analysis. All four collaborative groups were recorded successfully on this occasion.

7.3.2 Topic 2

The topic studied was "the Earth in space", and was related to the Programme of Study for Physical processes (Great Britain, DES, 1991). The teacher expected the children to learn about: the planets and stars as major types of body in the universe; and the relative movements of the objects in the solar system. The terms to be mapped were chosen by the teacher:

universe; stars; planets; sun;

Earth; moon; satellite.

These terms were reproduced on a sheet of paper, which the children could cut up. On each occasion, a single map was made by each group, but each child had a copy of the words to help focus their thinking. In addition to the teacher and researcher, there was a non-teaching classroom assistant present for each session.

The first, pre-topic, map was made before starting work on the topic. Audio-tapes were made of each of five groups, and copies of these groups' concept maps were retained by the researcher for analysis. One of the recordings was incomplete, due to the children's having turned off the tape recorder and failing to set it back correctly. The post-topic session took place two weeks later, after work on the topic was completed. Two of the original groups had one member absent on this occasion, and another had two children present who were absent for the earlier session. Successful recordings were made for each of the groups, and copies of the completed concept maps were retained.

7.3.3 Topic 3

The third topic was "Sound and hearing", and was related to the Programmes of Study for both Physical processes and Life and living processes (Great Britain, DES, 1991). The work covered: the parts of the ear and their rôle in hearing; and the nature and qualities of sound waves. The terms to be mapped were chosen by the teacher. For the pre-topic session, these were:

sound waves; echo; loudness; ear drum; pitch;

decibels; vibrate; guitar string.

These were given to each of the children in the groups on a sheet of paper, which they could cut up. They were permitted to add words of their own. Each group produced a single concept map.

The first session took place at the time of a local inspection of the school, and there was an inspector present through the session, who also interacted with the children in the groups. Four out of the five groups were successfully recorded, the one failure due to battery failure a short way into the session. Copies of the completed maps were retained for analysis.

The post-topic session took place four weeks later. For this map, the teacher changed the terms to be used, as she had not covered all the work that she had intended. The words given to the children were:

sound waves; echo; ear drum; pitch;

vibrate; guitar string; cochlea.

On this occasion, there was a model ear with removable parts in the classroom, as well as a guitar that was normally in the room. Some of the children referred to these in the course of the session. Five groups were recorded successfully, but only four of the groups completed their concept maps. A post-topic concept map was not obtained for group 14 as they had committed very little to paper by the end of the session. Copies of the completed maps were retained.

7.3.4 Data Set for the Main Study

The data collected for the main study have been described in previous sections. In Table 7.1, the main data set is summarized, and the relationship between the different elements shown. Where the data collected were incomplete, there is nothing to suggest that these were anything other than random losses. Data from the partial recordings for groups 6 and 12 were available to contribute illustrative material, but were not sufficiently representative of entire discussions to contribute to quantitative comparisons.

Table 7.1: Data Collected

Topic	Group	Pre-topic Session		Post-Topic Session	
		Map	Tape	Map	Tape
1	1	✓	×	✓	✓
	2	✓	×	✓	✓
	3	✓	✓	✓	✓
	4	✓	✓	✓	✓
2	5	✓	✓	✓	✓
	6	✓	×	✓	✓
	7	✓	✓	✓	✓
	8	✓	✓	✓	✓
	9	✓	✓	✓	✓
3	10	✓	✓	✓	✓
	11	✓	✓	✓	✓
	12	✓	×	✓	✓
	13	✓	✓	✓	✓
	14	✓	✓	×	✓

Key: ✓ = data collected × = data not collected

8

DATA ANALYSIS METHODOLOGY

8.1 Basis for an Analysis of Pupil Talk

In Chapter 2, it was discussed how meanings are constituted and maintained in social interaction. It was necessary for the purposes of the present study to derive methods of analysis that enable these theoretical insights to be applied in investigating the patterns of interaction that take place during group concept mapping. In the present chapter, it will be shown how existing theory was drawn upon in formulating a working methodology, matched to the specific needs of the study.

In Chapter 7, it was proposed to develop a classification of the data through a process of iteration between prior conceptions and encounters with the data. The prior conceptions brought by the researcher to this study were of two kinds. Firstly, there were those notions about learning and interaction that have been discussed in previous chapters, and that form the focus of the investigation. These had an immediate influence on what data were collected, as well as on how those data were classified. Then secondly there were intuitions about the talk generated in discussion groups that arose from the researcher's being a native member of the language community that sustains the possibility of this type of discourse (Stubbs, 1983). On hearing what participants say, one by and large knows what they mean. Stubbs sees membership of the language community as a resource, rather than an obstacle to understanding, and this view is consistent with the hermeneutic underpinnings of this analysis. However, in exploiting this resource, he recommends using an "estrangement device" as a way to "step back" from the immediacy of perception.

The "estrangement device" used in this part of the research was an evolving system of categories imposed on the data. Because the system was regarded as provisional, the possibility was created of a dialogical relationship between theory and data. Having formed initial categories, these were applied to the data, and the resulting coding patterns were compared to an intuitive understanding of what the participants were trying to achieve through their talk. Subsequently, the categories evolved through an iterative process of modification and re-application until they were capable of accounting for an acceptable proportion of the data. In doing so, they came to reflect more of the meaning of whole exchanges. But at the same time, and through the same process working in reverse, as it were, the analyst's views about the meaning of the exchanges also developed. The final analysis system must therefore be seen as neither directly derivable from theory, nor embedded in the data and requiring only to be isolated.

8.2 Development of the Analytic System

A potential pitfall with qualitative analyses of the kind suggested above is that much of the interpretation is itself a private process undertaken by the researcher, and often must remain so, as there is nobody else sufficiently "close" to the data to be able to carry out this interpretation. Hence the insights developed may be deemed subjective, and validity called into question.

Concerns such as these have prompted Constas (1992) to argue for all aspects of a qualitative analysis to be open to public inspection. Whilst this cannot eliminate biased accounts, it goes some way towards encouraging a rigorous approach to improving validity and reliability, and enables the reader more readily to detect where this has failed. Accordingly, it is proposed to document systematically what prior assumptions formed the starting point for this analysis, the stages through which the analysis framework subsequently developed from these assumptions, and the measures taken to verify the reasonableness of the categories developed. Constas suggests tabulating this information, but a narrative account will be adopted here as being more informative, and more in keeping with the iterative nature of the present analysis.

8.2.1 Overview of the Development Process

The first step was to sketch out a set of preliminary categories. These enabled an initial exploration of the nature of the language used in the group tasks. This preliminary analysis system subsequently evolved through repeated comparison with the data gathered, and through further consideration of relevant literature. The stages through which the development passed are approximately as follows (though it must be recognized that this is a *post hoc* account, and it is not necessarily possible to ascribe a strict chronology to these "stages").

First, categories were derived selectively (but more or less directly) from existing relevant literature. These categories tended to be highly analytical, and in practice had the effect of over-emphasizing the linguistic structure of the discourse at the expense of the themes developed within it. However, this is in keeping with Stubbs' (1981) insistence that linguistic structures be recognized in any analysis of classroom talk, and in due course it facilitated progress towards a more revealing categorization.

Jacob (1987) distinguishes between operationally defined categories, which "are prescriptions of what to see", and the symbolic interactionists' "sensitizing" constructs, which "suggest directions along which to look" (p.30). With each subsequent revision, the system of categories was formulated in a fairly loose way at first, and was then applied repeatedly to the data. In doing so, it became clearer how the system could be improved through restructuring, deleting, modifying or adding categories, and ultimately how these categories should be reified for use in the main phase of the research. Since the nature of the categories is bound up with the way the category system is structured, their developmental history will be discussed below as part of the description of the system as finalized.

8.2.2 Data Theory: Introduction

In conjunction with an examination of some of the early recordings and a logical review of the possibilities, the initial set of categories was developed with reference to several strands of existing literature. These were:

- existing category systems used in research on group work (Barnes & Todd, 1977; Webb, 1982a, 1982b, 1989; Gilbert & Pope, 1986; Kempa & Ayob 1991; Kempa, 1993; Roth & Roychoudhury, 1993);
- linguistically orientated theory (Halliday, 1973; Sinclair & Coulthard, 1975; Barnes & Todd, 1977, 1981; Stubbs, 1981, 1983; Winograd, 1985);
- philosophical perspectives (Wittgenstein, 1967; Searle, 1969; Winograd 1985).

A review of existing research on group work formed the starting point in the search for a suitable classification. However, in determining the extent of its relevance it was important to examine some alternative perspectives on language that have a bearing on the analysis. For Stubbs (1981), discourse is "a highly patterned, rule-governed activity describable in terms of several interrelated ranks of description" (p.126). Different features of language reflect linguistic choices made at different structural levels (see also Halliday, 1973). For example, the specific sequence of sounds a speaker produces is fixed, once a decision has been made at the level of the words to use, and this in turn is decided, with reference to situationally specific rules, in accordance with the purpose for which those words are to be uttered. Stubbs therefore argues for a principled selection of the level at which linguistic features are analysed: we should not, in the previous example, look upon phonological structure as evidence of thought processes. Different analysis systems operate at different levels of organisation in the data, and hence may be suited for one purpose but quite useless for others.

As discussed in Chapter 3, in Webb's (1982a) research on cooperative learning, pupil talk is analysed according to broad categories, such as ***giving explanations*** as opposed to ***giving short-answer feedback***. Webb's work (and also the other research reviewed by her; Webb, 1989) was focused on closed-outcome problems, such as mathematical textbook exercises. With this focus, it might be considered a relatively straightforward matter to distinguish between, say, giving the answer to a problem alone and giving "descriptions of how to solve the problem or part of it" (Webb, 1989, p.25). The content of the utterance would normally be sufficient to categorize it. Furthermore, both feedback and explanations were viewed as responses to an identified need for help. The scenario, in other words, was a simple problem-response situation in which an individual got into difficulty solving a problem and this difficulty was communicated to other group members, either through a clearly incorrect problem solution or through expressing the nature of the difficulty. It is not clear how this relates to situations in which ideas are to be explored and discussed, that is, to activities concerned with *meaning*.

In activities of the latter sort, it could not be assumed that individual utterances were the source of advances in understanding. The sharing of ideas need not be only in response to mistakes or difficulties, and the accumulation of meaning could well take place over a large number of moves in a conversation, as threads in the discourse emerge, are set aside and then taken up again later. Along this trail of meaning, it is possible that ostensibly the same utterance could have different rôles in different stages or different contexts. The content of the utterance does not suffice in classifying it. A shift of perspective is therefore required, reflecting the distinction between two interrelated aspects of meaning: what *words* mean and what *people* mean. Each utterance needs to be considered and evaluated in its specific context, in relation to others preceding or succeeding it (but not necessarily directly preceding or succeeding it).

These difficulties militate against the direct transfer of Webb's methodology to this present research. However, Webb contributes the important idea that there are *levels of usefulness*, with some kinds of utterance being of more value than others in making progress with a task.

Gilbert's & Pope's (1986) research addresses more directly the processes involved in meaning change, and their approach would seem to offer a more promising way forward. Their set of categories was developed to analyse group discussions about single scientific constructs, and the resulting scheme blends together aspects of content and function of an utterance. Thus there are categories such as ***challenge to a conception*** and ***defence of a conception***. Some of these categories are relevant to the purposes of this present study, and will therefore be discussed further.

Roth & Roychoudhury (1993) took an overtly "anthropological" approach to examining the protocol transcripts from collaborative concept mapping sessions. They claim that their goal was "to develop categories from the data" (p.508), and that three major processes "emerged": ***collaborative construction of propositions, adversarial exchanges*** and the ***formation of temporary alliances***. However, the extent to which these categories actually did fall out of the data, as opposed to their being pushed out, should be viewed with at least a degree of scepticism. They appear to replicate certain categories of interaction associated with communication in scientific communities, as previously established by sociologists of science such as Knorr-Cetina (1981). (The authors do not make this last point until discussing their findings at the end of their paper.) As with Gilbert's & Pope's (op cit.) scheme, these categories feature elements of content and function, but by way of contrast are applicable to segments of discourse rather than individual utterances. They therefore exemplify a more holistic analysis of discourse operating at levels of organization beyond individual utterances. Some attention to larger units of talk would seem to be essential in meeting the needs of this present research.

When examining in detail a transcript such as those generated in this research, the apparently high ambiguity of much of what people say soon becomes apparent. Yet it is also clear that the participants, by and large, have little difficulty in making an adequate interpretation. Moreover, the listener too has little difficulty. The difficulty arises only when it is necessary to make some sort of classification of how particular utterances contribute to the whole.

8.2.3 Towards a Solution: Speech Act Theory

The primary concern of the research is with the construction of *meaning* (as usage, à la Wittgenstein). The meaning of an utterance is dependent on its context, and is underdetermined by the surface features of what is said and its content (Stubbs, 1981, 1983). Consider, for example, the utterance "Who was the last one in?". This has the form of a question, and appears to be asking for information. It is possible to envisage situations in which this would indeed be its function (perhaps a race to dive into a swimming pool would be one). Alternatively, uttered by a teacher waiting to start a lesson who notices that the door is still open, it might serve as a command to close the door. This example also illustrates the possibility of multiple simultaneous functions of utterances: in the case of the teacher's usage, the utterance would most likely serve as a reprimand, as well as a directive.

The theory of *speech acts* has proved seminal in the philosophy of language and in sociolinguistics. This theory, which develops the Wittgensteinian "meaning-as-use" view, was originated by J L Austin (Searle, 1969; Stubbs, 1983). It accepts that the surface features of an utterance form only one aspect of its meaning in a given situation, and adds the insight that utterances are *actions* in that situation. Thus we do not simply make statements for their own sake, but have some purpose for doing so. A substantial aspect of the meaning of any utterance is what it counts as in the particular context in which it is uttered; whether it is a threat, a promise, a greeting or so on. Naturally, the form and content of what is said are constrained by the kind of act intended, but are not fully determined by those acts. Stubbs (1983) illustrates this with reference to the act of promising. "I promise I will be there", "I will be there without fail" and "I will be there ... Without fail? ... Yes" (the latter a cooperatively constructed act) are interchangeable ways of making a promise, though only one contains the giveaway "performative" verb, to promise. The term "illocutionary act" (or "illocutionary force") is often used to refer to the function performed in and by an utterance (Searle 1969). From an artificial intelligence perspective, Winograd (1985) likens speech acts to commands to run programs: an informative act is interpreted as a command to add information to a database.

In the example given earlier ("Who was the last one in?"), the illocutionary force is in the one case an elicitation, and in the other a directive (and probably a reprimand as well). One reason that the same surface form of words can have different functions in different utterances is that the intention of the speaker differs in each case. Doubtless there are subtle messages in intonation which help to disclose these intentions, but the listener is also dependent on an appraisal of the context in order to come to the right conclusions about what is meant. In the present example, relevant shared knowledge that would be drawn upon to determine the force of the utterance would be that it is customary for the last person arriving at the lesson to shut the door (and also, possibly, that teachers are given to sarcasm). Barnes & Todd (1981) refer to this background knowledge as the "frames" within which the interaction takes place.

The preceding discussion confirms there can be no direct way of classifying an utterance on the basis of syntax and propositional content. One must, in addition, enter the appropriate frame with the participants and determine what social acts they are performing. Entering the participants' frame may never be entirely possible for an observer, but by being present when the concept mapping sessions take place, and by listening to the tapes in conjunction with reading the transcripts (and hence benefiting from every nuance of tone and timing), it is possible to approach this ideal.

8.2.4 Discourse Structure

Within a frame, the nature of each linguistic act constrains what is permissible in response. The constraining relationship is twofold. Firstly, there are restrictions on what kind of act follows what. For example, a genuine request (whatever form it takes) predicts an answer, and this understanding can be used to tap intuitions about discourse in order to achieve a classification. So if a person says "I'd like to know when the next train is", our linguistic instinct suggests that the appropriate and expected response is to supply the time of the train, rather than to say "Thank you for telling me that", which would be classed as deviant. Intuition classifies the initiating act as making a request rather than as informing.

Use of the terms "instinct" and "intuition" might be considered to signal an undesirable element of subjectivity, so it is important to be clear about how they are being used here. What actually takes place when a listener makes judgements of this kind is the application of tacit knowledge of the workings of language generally, and spoken English in particular. This knowledge is built up over the course of a speaker's lifetime. It is not possible for the speaker to specify exactly what rules are applied in using the language, but this is not sufficient to deny that such rules exist. That the person succeeds (at least most of the time) in communicating with other speakers of the language indicates there must be shared rules. Searle (1969) refers to these as the "constitutive" rules that make the activity of language use possible; without such rules we would not be using *language*, whatever else we might be doing. So when a speaker instinctively recognizes what acts are taking place, it is because she or he can apply a (well-corroborated) theory. This is one way, then, in which prior theory determines how utterances are to be classified.

Various writers have shown that much discourse can be characterized by a generalized ***Initiation/Response/Feedback*** structure (Stubbs, 1983). An important early attempt at an analysis of discourse within this three part structure is that of Sinclair & Coulthard (1975). However, their main concern was to elucidate how discourse in general is structured, rather than how the topic of discourse develops, a purpose which led them to seek out highly regularized contexts of language use. Typical of this type of exchange is the sequence of classroom talk in which a teacher asks a question (initiation), a pupil responds with an answer to the question, and the teacher acknowledges the answer, normally by evaluating it (feedback). In more complex conversation, a response may itself serve as another initiation, which precipitates a further response, and so on. Feedback may take place at various times in the exchange. Hence, the discourse may be extended over a considerable number of these three classes of act. Such a structure can be identified in the transcripts from the present research. However, because it is intended to be a very general description of discourse structure, classification of utterances as initiation, response or feedback is insufficient on its own for the purposes of this research. We lose sight, for example, of whether children are initiating a discussion relevant to the task, or a discussion of the previous evening's television programme. The level of organization addressed is not, therefore, optimal in the present case.

The second type of relationship present in discourse is the logical relationship between the subject-matter of an utterance and those that follow it. For the earlier example, a request about the time of a train is not (without deviance) to be followed by a statement about the weather. Such structuring is more overt and straightforward than linguistic rules, but with the caveat that some common understanding of the content is necessary for conversation to continue. Where such common understanding breaks down, a process of repair must take place, or the conversation will fail. This is a matter of considerable interest for the present study, as it is the participants' theories about the subject-matter that the activity of concept mapping is intended to affect.

In framing the present analysis system, then, it is important not to divorce the above two kinds of relationship. What is required is an appreciation of how each utterance responds to both the content *and* purpose of those preceding it.

8.2.5 Background to a Speech Act: Strategy

The point was made above that people generally do not utter statements for their own sake, but have some purpose in mind. That purpose is not simply to perform speech acts for their own sake. We do not make promises unless we have a reason for doing so, and hence there is a level of organization above that of individual speech acts which seeks to incorporate such acts into an overall strategy. Stubbs (1983) shows that we must posit the existence of some such underlying motivation to account for coherence in discourse that is not explicable by surface structure. These underlying strategies relate to socio-cognitive needs (for example, to the resolution of conflict).

Since a person may well be acting on more than one strategy simultaneously, this accounts for how a single utterance can carry out more than one illocutionary act. For the purposes of this present research, we are interested in strategies that relate to developing and negotiating ideas. Hence speech acts that realize these strategies need to be analysed at greater depth. We are interested, for example, in recognizing whether the participants in a discussion are merely proposing, one after another, completed propositions for which the intended response is that they should be included in the concept map, or whether what participants offer is interpreted as an invitation to respond by adding, altering or correcting what is said prior to its being written down.

8.2.6 Holistic Concerns

Thus far, the tendency has been to concentrate on highly analytical approaches to the discourse, that operate at the level of individual utterances (albeit that this is seen as being informed by their wider context). Exclusive attention to such an approach would be in danger of obscuring how ideas might be developed over the course of a large number of turns in the discussion. Clearly, if concept mapping is to result in any knowledge restructuring, the information that is being discussed is important. The accuracy and structure of the scientific information arising from the discourse must be assessed. However, to attempt this for each individual utterance would not result in useful information. Most utterances made in the course of such discussions do not make determinate statements about scientific content.

Curiously, this aspect of pupil discussions seems largely to have been neglected in previous research. Neither Webb (*op cit.*) nor Kempa (*op cit.*), for example, takes a longer view of what is achieved in large units of discourse. Gilbert's & Pope's (1986) analysis methodology differs from these in that it is designed to trace the "history" of a group's negotiation, rather than the proportions of different kinds of utterance. However, they address, not the correctness of the conceptions being applied in the discussion, but how those conceptions are received within the group. In doing so, they do not classify the "histories" themselves. Of all the research reviewed, Roth's & Roychoudhury's (1993) work alone serves to show how larger categories may be derived (the above reservations notwithstanding) from the interaction between prior sensitizing constructs and the data.

In summary, an analysis at the level of utterances is required, so as to give a picture of the overall level of "interactiveness" within a group session. But this may miss crucial events in the development of understanding simply because these events are not reducible to atomistic elements. (Analogously, we do not understand fully how a computer works solely by classifying and counting its micro-components). However, it may be that having classified the function of utterances, we are better able to discern the larger units of discourse of which they form part. This was the approach adopted for the present study, and in the following part of the chapter, it will be shown how the various categories took shape.

8.3 The Category System

In the next section, some basic constructs that give structure to the category system will be described. In subsequent sections, the categories will be defined and their origins discussed.

8.3.1 Preliminaries

The most delicate unit of analysis for the system is a ***discourse move***. The term "utterance" has been used freely in the preceding discussion, but is not necessarily the most helpful way of dividing up speech. "Utterance" is a convenient way of referring to a recognizable and temporally isolated "chunk" of speech. Often this coincides with a move. However, without the underlying construct "move", there is no easy way of making sense of the situation in which a person is interrupted in the middle of making a statement, breaks off, and then resumes again when the other speaker has finished.

The term "move" is taken from Sinclair & Coulthard (1975). A move is a complete contribution consisting of one or more speech acts. It is "what the speaker is doing" at a particular point in the discourse. By way of illustration, a hypothetical segment of discourse might start with "Right ... lets have universe at the top". The single utterance "right" is a particular kind of act termed a "marker", which alerts listeners to a coming topic (*ibid.*), in this case introducing a discussion about hierarchy. "What the speaker is doing" consists in introducing this idea, and the marker act is subservient to that end. Similarly, a speaker may ask a question and then immediately answer it, in order to make a proposal or supply information. The purpose of the question is then not to elicit information from someone else, but to encourage others to think about what the speaker is saying. This definition of a move does not depend upon its being uninterrupted.

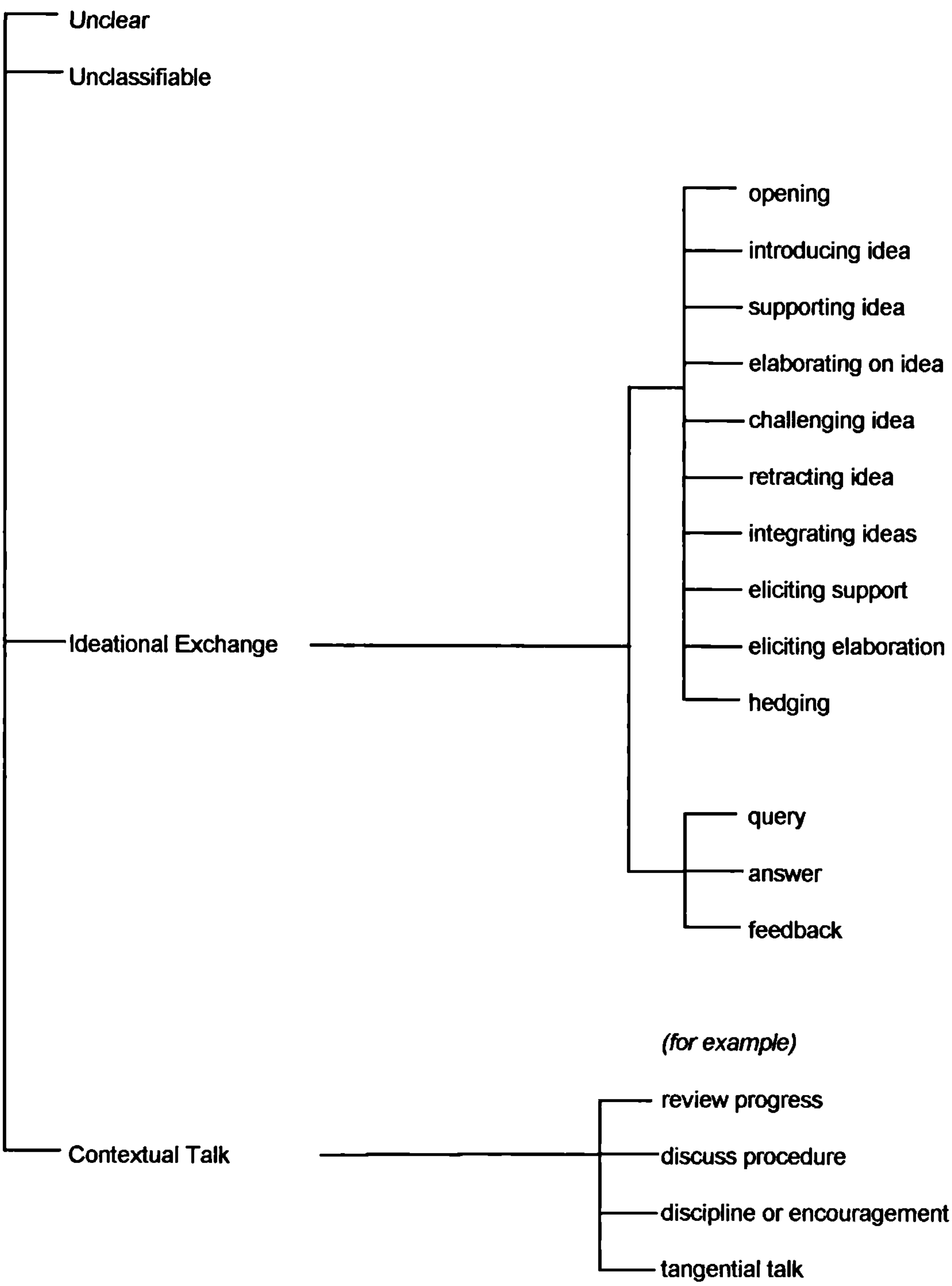
Moves are structural units linked by means of a higher level of organization that supplies coherence to the discourse. This is termed an ***exchange*** (Sinclair & Coulthard, 1975; Stubbs, 1983). It is at the exchange level that information is negotiated and shared meaning develops. According to Stubbs (1983):

> The basic notion is, then, that an exchange comprises an initiation and any contributions which tend to close that mini-topic: by completing a proposition ... ; by acknowledging it, and so on. (*ibid.*, p.135)

The purpose of analysing at the level of moves is to build up an overview of how exchanges in the discourse are structured and how ideas are developed within them. Moves are not, therefore, primarily of significance to the individuals, but to the group as a whole. Consequently no systematic analysis is made at the individual pupil level; indeed, in view of the stated position that moves only make sense in the context of exchanges as a whole, it is not clear how such an analysis could be made.

The system of analytic categories is set out in Figure 8.1. The category structure is presented in the form of a hierarchical tree of mutually exclusive choices, representing the range and configuration of the "meaning potential" open to each speaker (Halliday, 1973).

Figure 8.1: Analysis Network for Pupil Talk



8.3.2 The Categories: Introduction

Categories in the present system are not "classical" categories in the sense of having clear-cut boundaries defined by common properties (Lakoff, 1987). Rather, like Wittgenstein's (1967) example of "games", they are characterised by "family resemblances", so that examples may resemble each other in varied ways. In what follows, the categories will therefore be described prototypically, in terms of idealized features, and it will be shown how paradigmatic exemplars map into these groups.

The analysis network in Figure 8.1 applies to all moves regardless of who makes them. (The teacher is at times a participant in the negotiation of meaning within the classroom, and it was therefore important that the present analysis be equally applicable to teacher moves.) The move "enters" the network at the left.

Logically, this first stage of the analysis concerns the theme underlying the move, that is, what kind of thing it is that the participants choose to talk about. This sets the context of the talk, that is, which language-game is being played. With some category systems, an early dichotomy is made between utterances that are related to completing the task in hand and those that are not (Kempa & Ayob, 1991). Initially, the present category system preserved this same distinction. However, this turned out not to be helpful. On listening to, and reflecting upon, the recordings made of the concept mapping sessions, it became clear that the distinction was simplistic. Cooperative tasks of this type do not entail continuous theoretical discourse for their successful completion. There are lulls in the discussion that accompany setting out or arranging paper slips, drawing and writing or waiting for others to catch up. In these situations at least, and probably as an ongoing part of the whole task, there seems to be a need for the social lubrication provided by a degree of casual chatter. Such chatter is often characterised by humour, and especially by the "in" jokes that seem to be important in maintaining classroom relationships (Walker & Adelman, 1976). It is consequently misleading to classify this automatically as task-unrelated, a category that carries the implicit label "unproductive". In contrast, much of the talk is about the rules and procedures of the task, and whilst this is "task related" has little to do with discussing scientific meanings.

A major concern of the present analysis is therefore to distinguish between exchanges about the meanings of the construct terms to be incorporated in the map, and other kinds of talk concerning the context in which those meanings are being applied. Hence, the network initially divides moves according to the kind of language-game to which they belong. Talk focused on the terms to be included is part of an ***ideational exchange*** (an exchange about ideas; the term "ideational" derives from Halliday's, 1973, work). Ideational exchanges are the centre of interest in the analysis, and the moves within them are subject to further detailed analysis to track how ideas are introduced and processed by the group. Consequently, what counts as belonging to an ideational exchange will become clearer as that further analysis is described.

The other main category of talk is ***contextual***. This is actually a cluster of overlapping kinds of talk, and is focused on such topics as the "nuts and bolts" of how physically to draw in the links on the map, the instructions and procedures to be followed and other aspects to do with the process of making the map and the setting in which it is being constructed. Although further analysis of such exchanges is not required, examples of subcategories are discussed below in order to clarify what is included and, by contrast, to reinforce what is meant by an ideational exchange.

Unclear utterances are utterances that cannot be transcribed with sufficient accuracy to enable categorisation, and are filtered off at this initial stage. These are the equivalent of "missing data", and are an indication of the adequacy or otherwise of the recording for the group in question. However, even when some of the words in an utterance are inaudible, there are sometimes sufficient cues remaining to make a classification possible, and in this case the move as a whole would not be classified as unclear. In addition to unclear utterances, segments of discourse are from time to time omitted from the analysis because it is evident that no useful information would result from examining them. Such might occur, for example, when a group breaks off a task to locate pencils or other equipment, or when the task has changed from constructing a concept map to transcribing it solely to produce a neater copy. These omissions are justified on the grounds that the discourse concerned is not accompanying the original task that is the focus of the investigation.

Unclassifiable utterances are those for which an adequate transcription can be made, but of which the meaning remains obscure. Such may arise, for example, if a significant part of the meaning is communicated non-verbally.

8.3.3 Contextual Talk

The boundaries between some of the categories in this section of the network are not distinct, and consequently the categories tend to blur into one another. However, as no analysis hangs on these distinctions, this is unimportant. What is important is the difference between contextual and ideational exchanges, and the categories below represent the main kinds of contextual talk, in order to make this distinction clearer.

Review Progress

From time to time, some groups feel the need to discuss the status of the task. Examples include how much time is left and how many terms still need to be incorporated, as well as comments about what has been achieved. Although such talk may refer to the relationships between terms, it is not concerned with changing *meanings* in any way. For example, a participant in the activity may read back what has just been written. Exchanges that review progress may be of importance to the smooth completion of the task, though that is beyond the scope of this analysis.

Discuss Procedure

Concept mapping is an activity that has certain conventions and procedures to follow, and in addition to these, teachers set up expectations about how tasks are to be undertaken in the particular context of their classroom. This category includes discussion of such conventions as the direction in which arrows should be drawn (excepting where the intention is to clarify the meaning of a link connected by such an arrow). It might also include talk about particular instructions the children have been given, such as whether or when they are allowed to glue down the paper slips. Talk that serves to rehearse the words to be written on the map is placed in this class: although primarily self-directed, this talk nevertheless makes public an aspect of private thought, thereby rendering it available for further discussion if appropriate.

Discipline or Encouragement

This category covers talk intended to focus attention on the task in hand, to encourage participation, to object to inappropriate behaviour, and so on.

Tangential

This category covers all talk that does not contribute directly to advancing the task, referred to above (misleadingly) as "task-unrelated". It includes the general "banter" associated with ongoing classroom relationships.

Marker

These are utterances such as "right", "okay" and so on, which are part of the means used, for example, to preserve the flow of the discourse. Often, markers do not serve as moves in their own right, but form just one act of a wider move and are therefore not classified separately (see 8.3.1 above).

8.3.4 Ideational Exchanges

Ideational exchanges concern the meaning of the construct terms used in the concept mapping task. Aspects of meaning that may be addressed are the hierarchical nature of the relationship between constructs, and the propositional content of those relationships. Roth & Roychoudhury (1993) analysed students' discussion whilst making concept maps into talk about different levels of hierarchy and talk that "expressed a clear relationship" (p.511). Initially, this was attempted with the present data as well. However, acquaintance with the data soon demonstrated that talk can often range between these different aspects of meaning over the course of adjacent moves. To attempt to classify these different moves as belonging to distinct types of exchange would be to deny the underlying coherence evident in the discourse. It also transpired that relationships were often expressed in a way that was far from "clear", and that utterances range along a continuum from the vague (which might be attempting to describe either hierarchy or the linkage) to precise statements (which could confidently be classed as about one or the other). Indeterminacy was identified earlier (3.2.4) as providing for the growth of ideas in a social setting. Hence the way in which this kind of development takes place over a series of moves is of considerable interest in the present study, and it was important not to code it out of existence by trying to pin each individual utterance down as belonging to one type of exchange or another. Hence only one, broad type of exchange was postulated for this part of the network. Such an exchange might begin with deciding which term goes below a certain other term, and continue with a discussion about the wording to express the relationship between the two terms.

The categories of move in this part of the network are of great importance in the analysis, and hence it is necessary to document their development. The next few paragraphs show the genesis of those categories and trace their subsequent evolution.

Following previous literature on the structure of discourse (as discussed in 8.2.4), it is possible to identify in the data moves that have initiating, responding and feedback functions, but this in itself is not a sufficient classification for present purposes. As Stubbs (1983) has observed, any notion of negotiation of meaning is largely by-passed in the Sinclair & Coulthard system, which seeks instead to identify the linguistic devices that make the structure of discourse evident to the participants. The present analysis makes somewhat different demands.

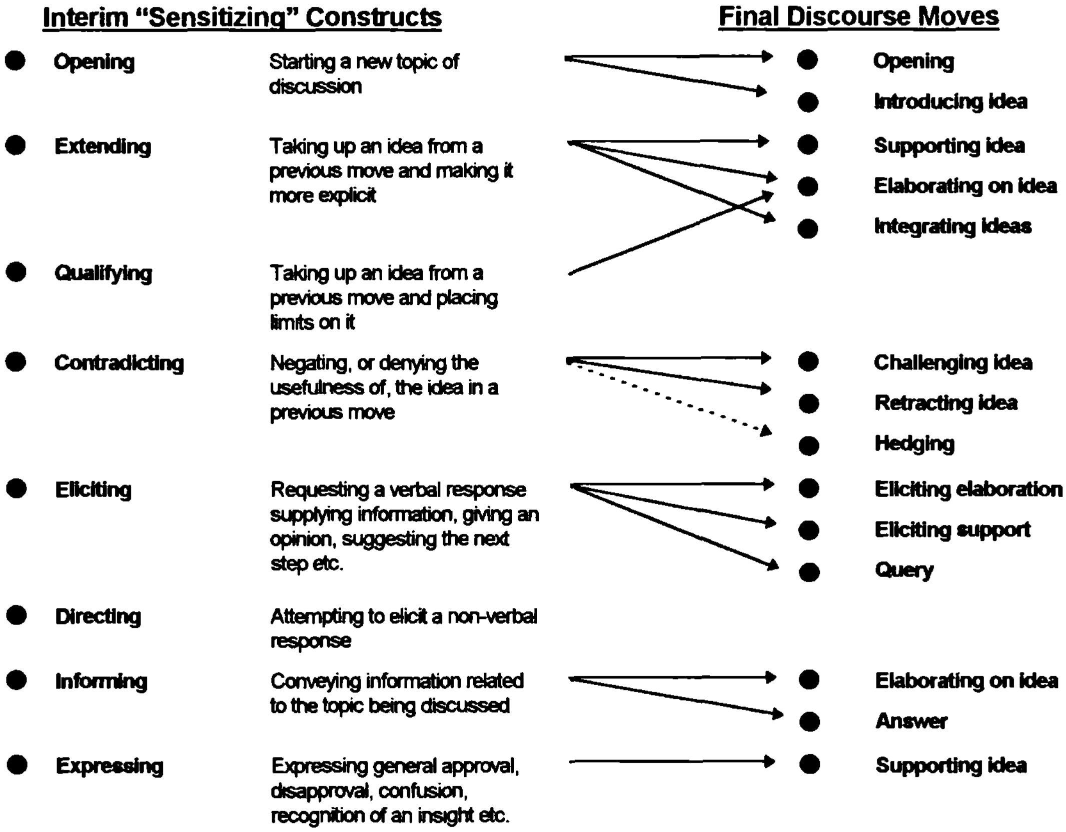
Barnes & Todd (1977) provide a more restricted, but contextually more relevant, analysis of pupil talk, developed for use with pupils undertaking collaborative problem-solving tasks. What Barnes & Todd tried to do was to trace how discourse is structured at different functional levels, and to identify relationships between the various levels. Thus at one level of analysis are types of "discourse move", akin to the categories of move explored by Sinclair & Coulthard, but with explicit acknowledgement that each such move is related to others not only by linguistic structures, but also logically, through the ideas expressed. This echoes the point, made earlier, that discourse is structured by both content and purpose. Barnes' & Todd's categories thus shaped an interim set of "sensitizing" constructs which, when applied to the data, guided the development of an appropriate set of categories for the present analysis. Figure 8.2 shows the relationship between Barnes' & Todd's categories and the interim classification used in this study.

Figure 8.2: Interim Categories of Pupil Talk

Barnes & Todd (1977) Discourse Moves		Interim "Sensitizing" Constructs	
Initiating	Starting a new topic of discussion	**Opening**	Starting a new topic of discussion
Extending	Taking up the initial idea of another	**Extending**	Taking up an idea from a previous move and making it more explicit
Qualifying	Placing limits on the idea of another	**Qualifying**	Taking up an idea from a previous move and placing limits on it
Contradicting	Denying the validity of the idea of another, proposing contradictory requirements etc.	**Contradicting**	Negating, or denying the usefulness of, the idea in a previous move
Eliciting	Requests to continue, expand on a point, give support or supply information etc.	**Eliciting**	Requesting a verbal response supplying information, giving an opinion, suggesting the next step etc.
		Directing	Attempting to elicit a non-verbal response
Responding	Continuing, expanding, giving support or supplying information etc.	**Informing**	Conveying information related to the topic being discussed
Accepting		**Expressing**	Expressing general approval, disapproval, confusion, recognition of an insight etc.

Note: bracketed categories should be considered together when making comparisons between the systems

Figure 8.3: Relationship of Interim to Final Categories of Pupil Talk



Applying these interim categories led to the realization that some important distinctions were obscured by the resultant classification, whilst other distinctions were insufficiently exact to be useful. In particular:

- the interim category of *opening* embraced two types of initiation move that needed to be distinguished;
- a large number of moves across all the transcripts neither added any new information to the discussion nor explicitly expressed approval, but still served mainly to support a previous suggestion or to keep it "ticking over". This class of utterance was not adequately represented in the system;
- there was a need to distinguish between substantive queries that led to an elaboration of meaning and those that served, for example, to confirm a previously stated formulation;

- there was a range of "backing down" moves that indicated a change of mind or a softening of attitude regarding a previous assertion. These were not distinguished in the initial categories.

These needs were addressed through a restructuring of the category system. Figure 8.3 illustrates the mapping between the interim and final categories, and in the following paragraphs, the finalized categories of move in ideational exchanges will be explicated. In the illustrative examples, "X", "Y" and so on are used to stand for construct terms.

Opening

This is always an initiating move. It is "opening" in the sense of "opening up" the discussion. It is thus open-ended, but having the function of inviting a more specific discussion of meaning. Normally (but not necessarily) expressed in the form of a question, examples include utterances such as:

- "Which one shall we do next?";
- "What's a good word to join on to X?";
- "Let's think of one to put at the top".

Introducing idea

This is also an initiating move, except when it follows an *opening*, in which case it may be considered both a responding and initiating move. It serves to introduce new specific ideational content to be discussed, however vague that content may be to begin with. It may be realized by a statement or a question. Examples would include:

- "How about X?";
- "Shall we put X at the top?";
- "X could join on to Y".

Once an idea has been introduced, the participants in the discussion have the choice of supporting or rejecting it, and within these choices, of modifying the idea if this appears fruitful.

Supporting idea

This is necessarily a responding move, which may be realized in a variety of ways. It is an expression of general approval of a preceding move, or a way of maintaining it as a topic of discussion. An important characteristic of supporting moves is that they introduce no new propositional content into the discussion. They may support an *introducing* move, or one of the kinds yet to be defined. Examples include:

- brief utterances such as "Mmm", "Yeah" and so on;
- repetition or paraphrase of all or part of a preceding utterance;
- non-elaborative responses to challenges such as "I know because it's true".

Elaborating on idea

This is also necessarily a responding move, though it may also serve to initiate further moves. It is broadly supportive of the move that it is a response to, but adds new propositional content to the discussion. There is a wide range of possible examples, and these may be realized as questions as well as statements. This type of move allows vaguely expressed ideas to be filled out or modified, but it is does not explicitly reject previous ideas unless it is the elaboration of a *challenge* (see below). In giving examples, it is necessary to show how an elaboration relates to a previous move, thus generic and specific examples are:

- "X is a kind of Y" elaborates upon "X goes with/joins on to Y";
- "Green plants need sunlight" elaborates upon "plants need sunlight";
- "Plants need sunlight to make food" elaborates upon "plants need sunlight";
- "Moons are satellites" elaborates upon "not all satellites are metal".
- "Are all Xs like that?" (- eliciting elaboration) "No" (- elaborating)

It is not necessary for an elaboration to be correct to be included in this category.

Challenging idea

This is a responding move (there must be something to challenge), but it is very often also an initiating move. It is an explicit or clearly implied rejection of the idea expressed in a preceding move. Subsequent moves that support or elaborate upon this rejection are not classed as challenges, but as supporting moves, elaborating moves, and so on. Introducing an alternative idea or elaboration does not count as a challenge unless accompanied by an element of rejection. Thus the distinction between this and other kinds of response may be quite fine. Examples include:

- "No", perhaps followed by an alternative proposal;
- "What?", or the repetition of a previous word or phrase, with intonation indicating disbelief, amazement or similar;
- contradiction, of the kind " X isn't anything to do with Y" or "X isn't a Y, it's a Z";
- a suggestion that there are insufficient grounds for an assertion: "Are you sure?".

Retracting idea

This is a move in which a speaker backs down on an idea previously expressed, often in the face of a *challenge*, but sometimes after reconsidering what has been or is being said. The speaker may or may not then immediately introduce an alternative idea, or propose a different elaboration. If either of the latter is the case, this is coded as a two-move utterance, that is a retraction followed by an *introduction* or an *elaboration*. Examples include any acknowledgement of the incorrectness of what has been said, often including the word "No", as:

- "X is a Y .... oh ... no it isn't";
- "Alright then, what about ....";
- "No, you're right".

Integrating ideas

In this move, two conflicting ideas within an exchange are reconciled, or a link is made between two or more ideational exchanges. It is a specialized form of elaborating move. Examples include:

- "If we put X in, we can link Y and Z to it" (where Y and/or Z is the subject of preceding discussion);
- "X is a Y, but it's a Z as well";
- "X is the most important idea, because all the others join up to it".

Eliciting support

Although support is often forthcoming, sometimes it is explicitly invited. Typically, this move is of the form:

- "Do you agree with that?"

It does not apply to the "tag" question ("isn't it"), which is typically a rhetorical device subservient to another act.

Eliciting elaboration

When participants in the discussion feel the need for more information about an idea under consideration, this move can be used to attempt to elicit an *elaborating* move. It is part of a strategy to add to the information available to the discussants, and is not the move used simply to confirm or make clear what has already been said or implied. Some examples would be:

- "Why do you think we should join X to Y?"
- "Is X a kind of a Y?"
- "I don't know what X means"
- "Where's the best place to link X to?"

Hedging

This move is a noncommittal response to a *request* or *challenge*, or a softening of a statement, or a supporting move expressed with reservation. It lies conceptually between a *supporting* move and a *retracting* move. Examples are often of the form:

- "Well, sort of"
- "Maybe"
- "Hmm" (with a falling then rising tone)

Query loop

This set of closely related moves has the function of clarifying or confirming a meaning or a decision. It consists of an initiating move, a responding move and an optional feedback move. These are termed ***Query, Answer*** and ***Feedback***. The ***Query*** is not intended to elicit new information, and often is simply a request to repeat something. It is thus ideational in only a marginal sense. Normally, the query leads to an ***Answer*** that supplies the requisite confirmation, repetition or clarification of a previous utterance. However, it can also lead into an *elaboration*, should the person responding decide to add to the stock of information being considered. This effectively breaks the loop, and returns to the ideational exchange proper. The

Feedback move is an acknowledgement that the answer has been received. Feedback moves can occur in ideational exchanges other than query loops, though they are in practice rare, and often subordinate to another move. Examples of query loops are of the form:

- "What did you say joins on to X?" - "Y" - "Oh yeah"
- "So we put X is a Y?" (which has already been suggested) - "Yeah"
- "It is true that an X is a Y, isn't it?" - "Yeah"

In the latter two examples, if the response had not simply been "yeah", but an alternative new item of information, then the loop is broken and this would *not* count as an answering move as defined here.

An example transcript appears in Appendix E to illustrate the methods and constructs introduced in this chapter.

8.3.5 Dependability of the Classification

Reliability of the analysis network was estimated by having a second rater recode one whole transcript using the corresponding audio tape, selected at random. The rater was an educational researcher and postgraduate student with qualifications in psychology, who was not aware of the aims of the research. Training in the coding scheme consisted of studying a draft of the present chapter and working through the example transcript.

Decision consistency statistics were calculated between the two sets of codes (Subkoviak, 1980). These were: the proportion of agreement, P_o, and the kappa coefficient, κ (Cohen, 1960). Cohen's kappa is the proportion of exact agreement for nominal scale variables beyond that expected by chance. The assumption in calculating κ is that all misclassifications are equally unacceptable. This gave $P_o = 0.85$ and $\kappa = 0.82$, which were considered adequate for an analysis of this type. The largest area of discrepancy was for the second rater to allocate to the category contextual talk moves that were originally coded as either belonging to query loops (three instances) or as supporting moves in ideational exchanges (three instances). This would have a minimal effect on subsequent analyses.

Since this comparison indicates the extent to which the two raters interpreted the children's utterances in the same way, this also constitutes validity evidence for the coding scheme, and consequently for the further analyses carried out based on the coding.

8.4 Analysis of Concept Maps

The concept maps produced by the groups are an important piece of evidence regarding what the children accepted as the outcome of their discussion. Some means of analysing the data provided by these concept maps is therefore needed.

There has been a great deal of discussion over whether and how concept maps should be "scored" (see for example Malone & Dekkers, 1984; Novak & Gowin, 1984; Novak & Musonda, 1991; White & Gunstone, 1992). Much of this discussion centres on the extent to which certain features of concept maps should receive extra weighting, and results from their proposed use in assessment, for which a single score is often desirable.

White & Gunstone (*op cit.*) sensibly suggest that concept maps are best interpreted qualitatively, but that scoring may be appropriate for some purposes. The present research is concerned with growth of scientific meaning, and hence it is helpful to be able to compare the overall levels of understanding attained by the various groups, and to relate this to the quality of the group discussion. To achieve this (whilst accepting fully the comments above about scoring), some form of quantitative measure is convenient. Hence a simple quantification procedure was adopted.

In 2.2.4, it was argued that the meaning of a construct resides only in the way it is related to others. Hence, for any group learning about a given topic, growth in knowledge will be reflected by a change in the way they represent the connections between relevant constructs in the domain. In this sense, the unit of meaning will be one such relationship. Propositional relationships represented in concept maps have two aspects.

Firstly, and most straightforwardly, credit needs to be awarded for each labelled link shown on the map that corresponds with a scientifically acceptable proposition. All analytic scoring schemes for concept maps feature this form of coding, and awarding one mark per correct link is a generally accepted measure. Opinions diverge, though, about how to treat links that do not conform to scientifically acceptable propositions, with some sources recommending the deduction of marks for misconceptions (Novak & Musonda, *op cit.*). For the purposes of the present study, a definite misconception is a "negative result", and so it was important to code examples of these. In addition to clear misconceptions, children sometimes made links that were excessively vague, or were examples of "everyday" usage with no scientific relevance, such as "stars come out at night". These, too, needed coding separately, as they contribute nothing to the scientific meaning of the terms involved. Scores were therefore recorded for each of the above types of relationship (correct, vague or incorrect), and separate totals found for each category. These categories and the corresponding scoring system may therefore be summarized thus:

- Each link expressing a relationship between constructs that corresponds with a currently acceptable scientific relationship: +1 mark;
- Each link that does not express a clear relationship, or that expresses a relationship, meaningful in everyday terms, that does not correspond with any currently acceptable scientific relationship: 0 marks;
- Each link expressing a relationship that contradicts a currently acceptable scientific relationship: -1 mark.

Secondly, part of the meaning of some of the terms in a concept map is due to the hierarchical structure of their relationships to other terms, as argued in 2.1.4. Most scoring schemes for concept maps recognize this and award extra marks for hierarchy (though White & Gunstone, *op cit.*, express the reservation that not all domains have a clear hierarchical structure). Novak (Novak & Gowin, Novak & Musonda, *op cit.*) also stresses the importance of integrating links between different parts of a concept map. Superordinate constructs play an important rôle in this integration.

In constructing their concept maps, the children were encouraged to discuss and to represent hierarchical relationships, and the recordings confirmed that typically they did. However, the maps were not always drawn in a way that was sufficiently clear for the hierarchy recognized by the children to be identified. As a consequence, it was not possible to score reliably this aspect of the meaning represented in the maps.

To evaluate the reliability of this coding procedure, two measures were used, making use of a second coder. The second coder was an educational researcher with postgraduate qualifications in both science and education. Firstly, a random selection of 40 propositions represented in the concept maps across the three topics were coded by both coders. The two sets of ratings were crosstabulated to obtain the degree of agreement, and decision consistency statistics were calculated. This gave P_o = 0.93 and κ = 0.88, which was considered to be an acceptable level of reliability for this purpose.

To gauge score reliability for the substudy, all the relevant concept maps were rescored, and the two sets of scores correlated. The correlation coefficient used was the Pearson product-moment coefficient. This gave a correlation of 0.90 for the pre-topic maps and 0.99 for the post-topic maps, indicating very high agreement in the ranking of the scores from the two codings.

9

RESULTS: PHASE THREE

9.1 Overview

The purpose of this third phase of the research was to characterize the nature of the discussion taking place within collaborative concept mapping groups, to investigate whether the kind of discussion fostered was conducive to effective learning of science and to consider how concept mapping might be made most effective in this rôle. The research took place in three waves. Each wave took place in a different classroom, and spanned the period of a science teaching unit or "topic" in that classroom. Further details were given in Chapter 7.

The first evidence to be considered here concerns the substudy set up to compare individual and collaborative approaches to concept mapping in the classroom. This was undertaken to answer the question "Does collaborative or individual concept mapping better promote children's understanding of scientific meanings?". Although only a small element of the research, this formed an important basis for the wider study. The results indicated that it was indeed collaborative work that led to greater use of scientific language in the children's concept maps. This is the subject of 9.2. In subsequent parts, attention is turned to how collaboration achieves this effect. This entails determining the nature of the discourse that took place during the collaborative concept mapping sessions, and relating it to the scientific meanings inscribed in the children's concept maps. In the final parts of the chapter, aspects of the concept mapping activity that relate specifically to the learning of science are discussed.

9.2 Individual versus Collaborative Concept Mapping

The substudy on individual versus collaborative mapping was carried out during the first concept mapping topic: Habitats. The data collected consisted of the pupils' concept maps constructed during both pre-topic and post-topic mapping sessions. These will now be described.

9.2.1 The Substudy Data Set

Concept maps were constructed by all those participating. Within the collaborative groups, each child made his or her own concept map. These maps were collected and formed the first source of data. The concept maps were scored according to the criteria given in 8.4, and these scores were taken as indicating the appropriateness of the scientific language incorporated in the maps.

One child in the collaborative condition had to leave the post-topic session early, and so did not finish his concept map. Scores for the maps made by the children within any one group differed only slightly, and so it was reasonable to estimate a score for this missing map, based on the scores for the other members of the group. For this purpose, the mean score for the other three members was calculated and rounded down to an integer.

The results of the scoring are described in Table 9.1. The data suggest the following. Firstly, the level and distribution of the scores at the outset were similar for the two groups, but there was slightly greater variation amongst individual mappers.

Table 9.1: Descriptive Statistics for Concept Map Scores (Topic 1)

Pre-topic Concept Map Score

	Mean	95% Conf. Interval	St. Deviation
All Groups	8.8	±1.6	4.4
Individual (n=13)	9.4	±2.9	5.3
Collaborative (n=15)	8.3	±1.8	3.6

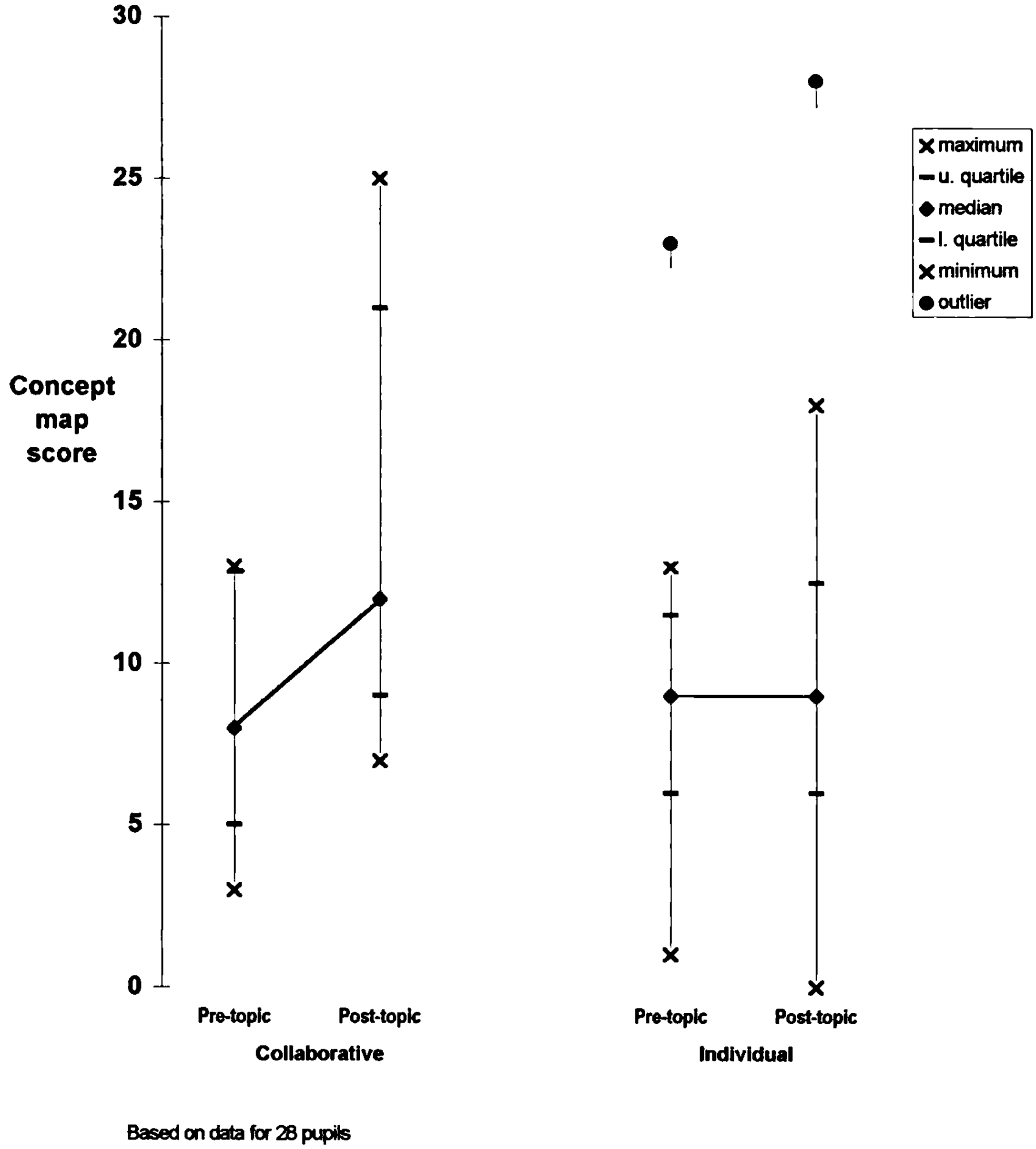
Post-topic Concept Map Score

	Mean	95% Conf. Interval	St. Deviation
All Groups	12.1	±2.5	6.6
Individual (n=13)	9.9	±3.8	7.0
Collaborative (n=15)	14.1	±3.0	5.9

The data were plotted to facilitate exploratory analysis, and the results are shown in Figure 9.1. This plot was helpful in revealing some important features of the data.

Firstly, there was a single outlying case in the individual mapping group for both pre- and post-topic sessions (the same individual in each session). This case would therefore contribute to the slightly higher mean score for the group in the first map completed, and to the greater variance. Setting this case aside, as is done pictorially in the figure, it appears that the two groups were quite similar in the level, midspread and range of scores for the first map. It appears, then, that there was little difference in the number of scientifically acceptable links displayed in the concept maps between the individual and the collaborative mappers. The plot also shows the nature of the change in scores from pre- to post-topic. There appears to have been little change in the scores attained by the individuals, with the main difference being an extension of the upward tail of the distribution. The middle part of the distribution appears to have changed very little. There was, however, an increase in the scores for the collaborative groups. This suggests that during the post-topic session, the collaborative mappers incorporated rather more appropriate linkages in their maps than the individuals. These findings were tested for statistical significance.

Figure 9.1: Schematic Plot of Concept Map Scores



9.2.2 Comparisons between Groups

For the purpose of hypothesis testing, the two groups were viewed as samples from two populations: an individual concept mapping population and a collaborative concept mapping population. The two populations were presumed to be equivalent in terms of uncontrolled variables at the outset, having been allocated through random selection within the class. The examination of the data reported above indicated that the distributions departed somewhat from normal, particularly in respect of the outlier. The sample sizes were not sufficient to ensure sample means would be normally distributed, so non-parametric statistical tests were used.

Firstly, the question was addressed: do collaborative concept mappers make scientifically more valid pre-topic maps than individual mappers? The Mann-Whitney *U*-test (Norusis, 1988) was applied to determine whether, for the populations represented by the two groups:

H_0: *the distributions of the pre-topic scores for individual and collaborative concept maps were equal.*

The α value for rejection was set at 0.05. The results of the test[15] showed that the null hypothesis should be retained. This supports the view that the two groups generated concept maps that were of equal quality at the outset of the study. On the basis of this evidence, it appears that there was no clear advantage in collaborating to produce the pre-topic concept maps. The next issue to be addressed was: do collaborative concept mappers make scientifically more valid post-topic maps than individual mappers?

Because any small differences in the children's grasp of scientific meanings resulting from the pre-topic session might have produced differential effects over the course of the topic, to maximize both the validity of the comparison and the power of the statistical test, an analysis of covariance procedure (Huitema, 1980) was used, adjusting for pre-topic score. Since a non-parametric test was required, Quade's rank analysis of covariance was used (*ibid.*). This procedure tests the null hypothesis that:

H_0: *the conditional post-topic score distributions were equal for the two populations.*

Quade's rank ANCOVA is reputedly the most powerful of the non-parametric ANCOVA techniques (*ibid.*). It consists of a parametric ANOVA carried out on the residualized ranks for the groups. Residuals were obtained by regressing post-topic deviation ranks on the pre-topic deviation ranks (the covariate) and calculating the difference between the predicted and observed post-topic rank scores.

Before applying the ANCOVA, it was necessary to check whether two major assumption were met: the relationship between pre- and post-topic scores should be monotonic, and the degree of monotonicity should be the

---

[15] $u = 89.0, p = 0.693$

same for both populations. A plot of the data confirmed the former (see Figure 9.2), whereas a homogeneity of regression slopes test on the rank scores was used to confirm the latter. The homogeneity test (see Huitema, 1980) uses multiple regression to examine whether the interaction between group and covariate accounts for a significant proportion of the variance. The obtained F did not exceed the critical value[16] and therefore was not statistically significant. It could be assumed that the regression slope was common to both groups, and the ANCOVA conditions were met.

The two-tailed probability was taken, as a difference in either direction was possible and was of interest to the research. The α value for rejection was set at 0.05. The results of the ANCOVA are shown in Table 9.2.

Table 9.2: Quade's ANCOVA on Post-topic Concept Map Scores

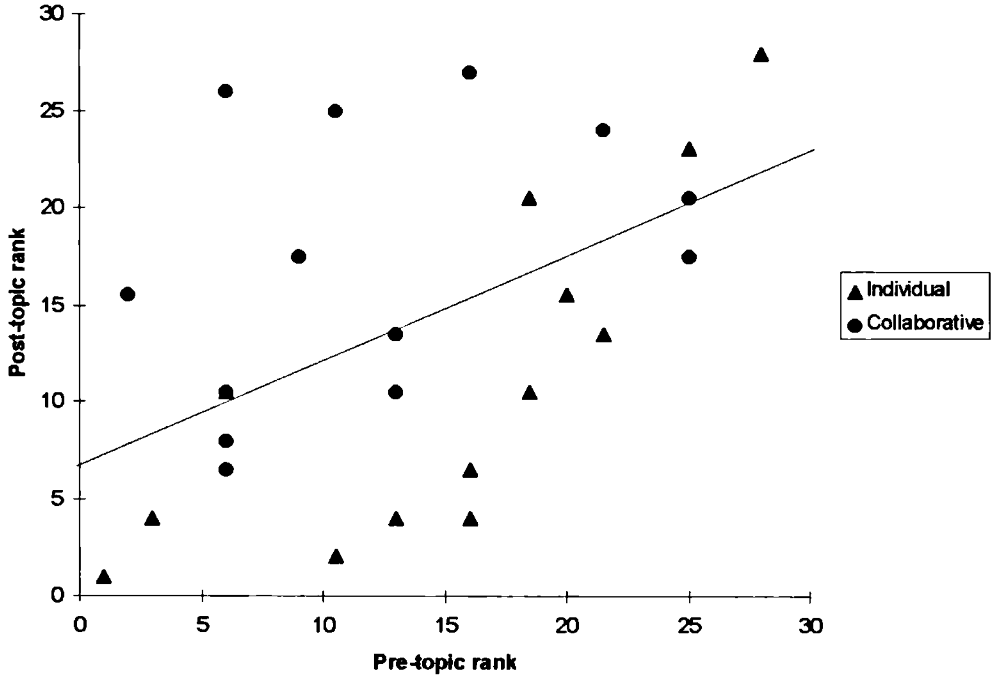
Source of Variation	Sum of Squares	DF	Mean Squares	F	Significance
Main Effects (Group)	359.043	1	359.043	9.848	.004
Explained	359.043	1	359.043	9.848	.004
Residual	947.963	26	36.460		
Total	1307.006	27	48.408		

Based on data for 28 pupils

The outcome of this test indicated that the null hypothesis should be rejected: the conditional score distributions were such that the collaborative concept mappers ranked higher overall than the individual mapping group. Figure 9.2 portrays the relationship between pre-topic rank score, post-topic rank score and group membership. For reference, the regression line is superimposed on the figure.

[16] $F\,(1, 24) = 3.37$

Figure 9.2: Regression Plot of Pre- and Post-topic Rank Scores by Group



Based on data for 28 pupils

The figure shows clearly that individual mappers tended to generate residuals below the regression line, while collaborative mappers featured mainly above it. It was concluded, on the basis of these various transformations of the data, that collaborative mappers made substantially more scientifically appropriate links in their post-topic maps than the individual mappers. This was the first piece of evidence that collaboration has an effect on the concept maps produced.

That a noticeable difference was found is in some ways surprising, since it might be expected that collaborative groups would spend longer coming to joint decisions about what links to make and how to describe them. Since the score is only one indicator of the nature of the final map, these findings warranted closer examination of the content of the two sets of maps.

The nature of the relationships generated between the construct terms did not differ greatly between the two groups. Both sets of pupils came up with similar links, and it is mainly in quantity that these differed in the post-topic session. There was a distinct tendency for the individual mappers to incorporate a greater proportion of vague links in their maps. Whereas these had largely disappeared from the collaboratively made post-topic maps, nearly all the individual post-topic maps included a number of links that were either unlabelled, or vaguely labelled, or did not convey a relationship that had any scientific meaning. The difference in the number of vague links in the maps between individual and collaborative mappers was tested for significance using the Mann-Whitney *U*-test. The difference was found not to be significant for the pre-topic maps[17], but was significant for the post-topic maps[18]. This was consonant with the analysis of the total map scores.

Both individuals and collaborators added construct terms of their own to the maps. The individuals contributed a greater variety of these, suggesting, perhaps, that the need to get additions ratified by the group tended to dissuade being adventurous. However, across both conditions, these extra links varied in their relevance to the topic, particularly in the pre-topic session. It is noteworthy, though, that only one of the collaborative groups introduced additional constructs in the post-topic session, and that these additions generated a well-integrated set of scientifically appropriate linkages concerning the rôle of plants and sunlight in producing the oxygen which animals breathe (see Figure 9.15 and also the related discussion). The additions made by the individual mappers in the post-topic session tended to be less integrated, although they too were mostly of some relevance to the topic.

Summary and conclusions

Taken together, these findings seem to suggest that constructing concept maps in collaborative groups affects the quality of the language used in the map, compared with constructing the maps individually. This effect was not detectable for the pre-topic mapping session, but emerged in the post-topic session, when there were, presumably, more links that the children could make as a result of knowing more about the topic. Whereas the collaborative mappers were able to incorporate more valid links in their concept maps at this stage, the individual mappers, in general, did not do so. The collaborative situation appears to have focused the children's thinking towards relevant connections. The following parts of this chapter will explore the nature of this apparent "group effect".

---

17 $u = 95.5, p = 0.92$

18 $u = 51.0, p = 0.03$

9.3 Analysis of Group Discussion

The principal thrust of this research is provided by the hypothesis that engaging in collaborative concept mapping will help children gain a facility with scientific meanings. In the preceding part of the chapter, the concept maps produced by collaborative groups were shown to be more oriented towards scientific meanings than those constructed individually.

These, however, were findings still very much in an experimental tradition that focuses on outcomes, and regards the "treatment" as an opaque process. It is the purpose of the following sections to open up the "black box" and to examine the workings of these processes. Quantified data are examined, in order to uncover patterns. The data are then subject to further qualitative analyses to provide interpretations of the processes underlying those patterns. The majority of the analysis consists of explicating what actions the participants were performing through the varying kinds of interaction present. What emerges from this analysis is a picture of a distinctive form of discourse, characterized by specific patterns of discourse move, which can be linked to the content of the pupils' concept maps. Together, these analyses will provide answers to the question "how was this outcome accomplished?" (Silverman, 1993, p.142). The results reveal that there is a connection between how ideas were discussed and negotiated within the groups and the quality of the scientific language incorporated in the finished concept maps.

The data set for this part of the research was described in Chapter 7, and summarized there in Table 7.1. The links made in the pupils' concept maps were classified using the procedure outlined in 8.4. Overall scores were not calculated because the quality of the individual propositions incorporated in the map was of interest, rather than the quality of the map as a whole.

Altogether, around 14 hours of audio-taped discussion were available across all the group sessions, comprising approximately 8000 discourse moves, and resulting in over 280 pages of transcripts. This constituted a substantial corpus of data. The 8000 moves were then classified using the analysis scheme described in Chapter 8.

As a result of this analysis, approximately 60 per cent of the moves were classified as consisting of contextual talk which was not of interest in this study. Less than five per cent were unclassifiable, either because they were insufficiently clear, or because they were unfinished, or it was otherwise impossible to determine the purpose of the move. This left just under 3000 moves classified as belonging to ideational exchanges, representing approximately 36 per cent of the data. The proportions of individual categories of discourse move within these exchanges are tabulated in Table 9.3.

Table 9.3: Frequencies of Categories of Discourse Move in Ideational Exchanges

Type of Move	Grp 1	Grp 2	Grp 3	Grp 4	Grp 5	Grp 6	Grp 7	Grp 8	Grp 9	Grp 10	Grp 11	Grp 12	Grp 13	Grp 14	All grps
opening	2	4	7	6	1	4	0	6	4	8	7	4	2	11	66 (2%)
introducing	20	35	57	31	13	20	22	20	20	26	28	15	27	32	366 (13%)
supporting	29	74	133	25	12	40	47	32	76	144	90	66	104	201	1073 (37%)
elaborating	21	47	46	18	12	26	18	13	27	114	56	31	71	74	574 (20%)
challenging	5	17	15	6	6	5	5	6	13	20	6	4	9	10	127 (4%)
retracting	1	3	2	30	2	3	2	1	3	1	3	1	4	5	61 (2%)
integrating	2	6	6	0	0	3	1	1	4	9	1	2	0	1	36 (1%)
eliciting support	0	0	1	0	0	0	1	1	3	5	11	1	6	10	39 (1%)
eliciting elaboration	5	5	2	1	4	8	3	3	10	41	27	14	29	24	176 (6%)
hedging	2	10	9	0	3	1	0	0	2	1	4	1	1	6	40 (1%)
query loop	12	21	37	24	10	16	23	15	70	34	35	19	26	24	366 (13%)
Totals	99	222	315	141	63	126	122	98	232	403	268	158	279	398	2924 (100%)

Although in only a fairly crude way, these data begin to reveal something of the kind of talk within the groups. The 36 per cent of the talk that was about the meaning of terms in the map may be compared with findings, for 6- to 7-year-old children, by Bennett *et al.* (1984). The nearest equivalent in their analysis was covered by the categories ***instructional input*** and ***sharing information: task-specific,*** which together accounted for just 18 per cent of pupil-pupil talk in both mathematics and language tasks. Against this baseline, the proportion here seems favourable. Turning to the proportions within the table, the number of introducing moves is fairly low compared with the remaining types. There were, on average, three supporting moves for each idea introduced. Similarly, each exchange featured, on average, just under two elaborative moves. Some six per cent of moves overall were prompts for further elaboration, while four per cent were direct challenges to an idea under discussion. All groups featured examples of both these latter kinds of discourse move. It appears, from this overview, that ideas were actually being developed over the course of several turns in the discussion, and not simply raised, accepted and transcribed onto the concept maps. However, it is necessary to look in more detail at the way the discussion proceeded in the groups to gain a clearer picture of the processes at work in producing a concept map collaboratively.

9.4 Ideational Exchanges and their Outcomes

The object of this second stage of the analysis was to identify the processes at work in the discourse, from the outset of the concept mapping task, when the children were presented with a set of construct terms, to the culmination of the activity in a completed concept map. This was necessary so that features of the discourse could be identified that were conducive to learning the language of science. It is only through identifying and understanding these processes and their relationship to learning that a picture of collaborative concept mapping can be developed that is of use to the teacher and that can indicate how practice might be improved.

Each ideational exchange consisted of a series of discourse moves, and it was expected that there would be patterns discernible in the sequencing and combination of such moves across the numerous exchanges in the data. Moreover, it was expected that specific kinds of exchange identified in this way would function differently in the overall task of compiling a concept map. The main guiding questions for this phase of the research were as stated in 5.3.7. They were operationalized as follows:

- *What are the main properties of the ideational exchanges taking place amongst the groups, and how do they vary?*
- *How do these properties of the discussion contribute to the construction of scientific meanings in the concept maps?*

Over the next few sections of this chapter, firstly, the properties of ideational exchanges will be described, and then the outcome from these exchanges will be analysed in terms of whether they led to scientifically appropriate meanings' being agreed upon and incorporated into the concept map.

9.4.1 Properties of Ideational Exchanges

Up to this point, the analysis has produced a fragmentation of the data into individual moves. But these moves are not functional in isolation. They must be seen as the building blocks of exchanges, and it is through exchanges that the business of compiling a concept map is accomplished. The next stage in the analysis is consequently a synthetic one, in which the way that individual moves are related to one another is examined. This process renders it possible to identify patterns in the structure of exchanges. The procedure has the ultimate goal of linking particular sequences of discourse move together with the resulting inscription in the concept map.

The main properties of ideational exchanges that were identified as pertinent to this study were:

- the length of the exchange;
- whether or not the idea introduced as the subject of the exchange was then further elaborated;
- whether any such further elaboration was carried out individually or collaboratively.

These properties will next be described and illustrated.

Length of exchange

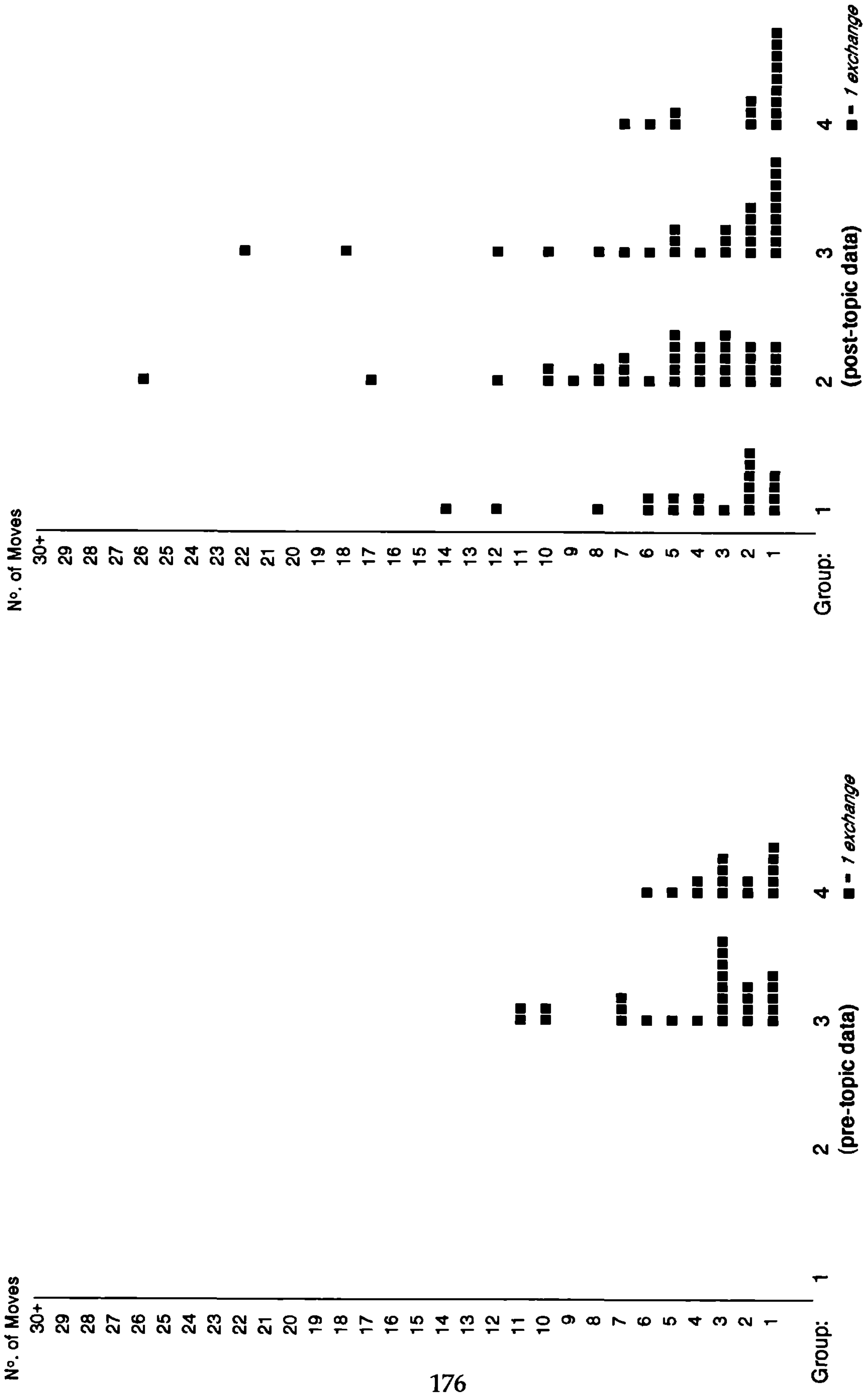
The length of an exchange is the number of moves making up the exchange. The first move in each exchange is the introducing move (opening moves being essentially preliminary to the commencement of the main exchange). The final move is when the idea is settled prior to its being incorporated in the concept map, or rejected as inappropriate. In cases where the introducing move is the only move in the "exchange", these are regarded as opening bids which the participants chose not to pursue further.

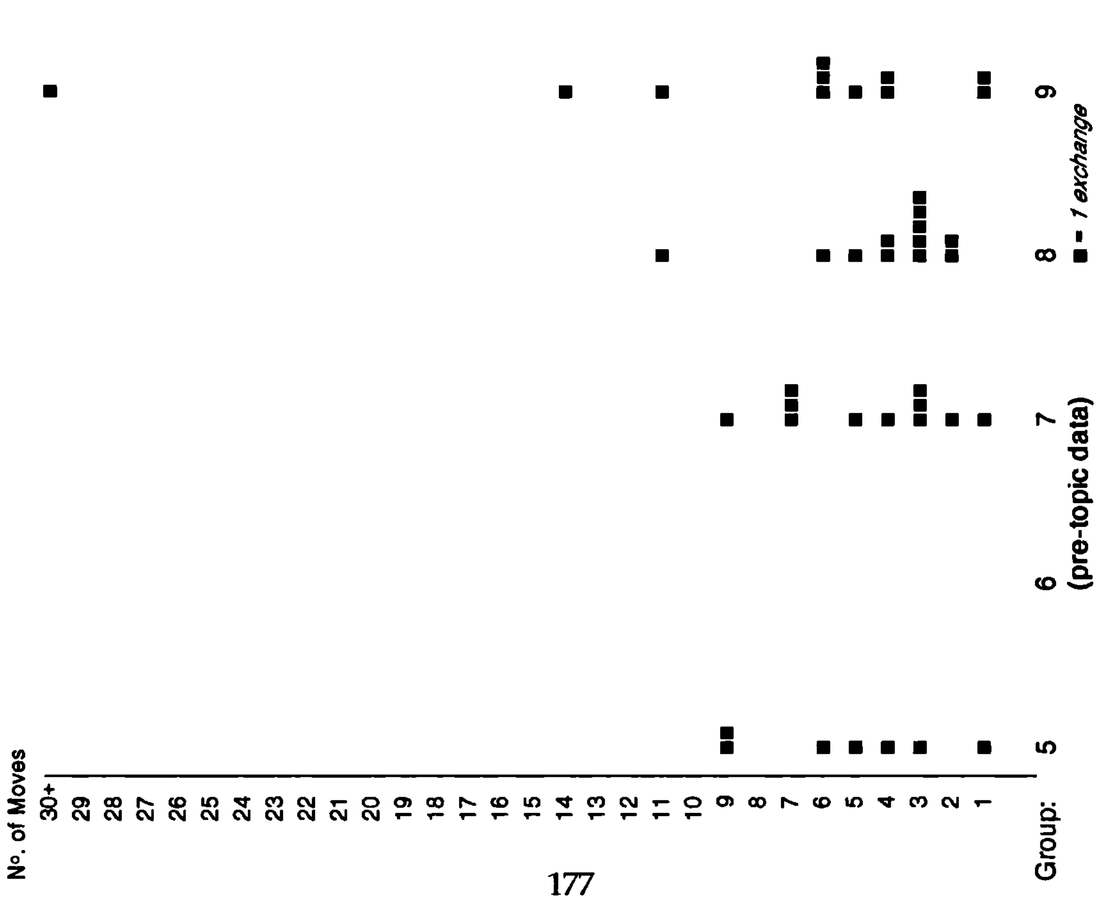
Counts of the number of moves per exchange were made for each group session. Figure 9.3 displays the number of moves per exchange for all groups for which complete data were available.

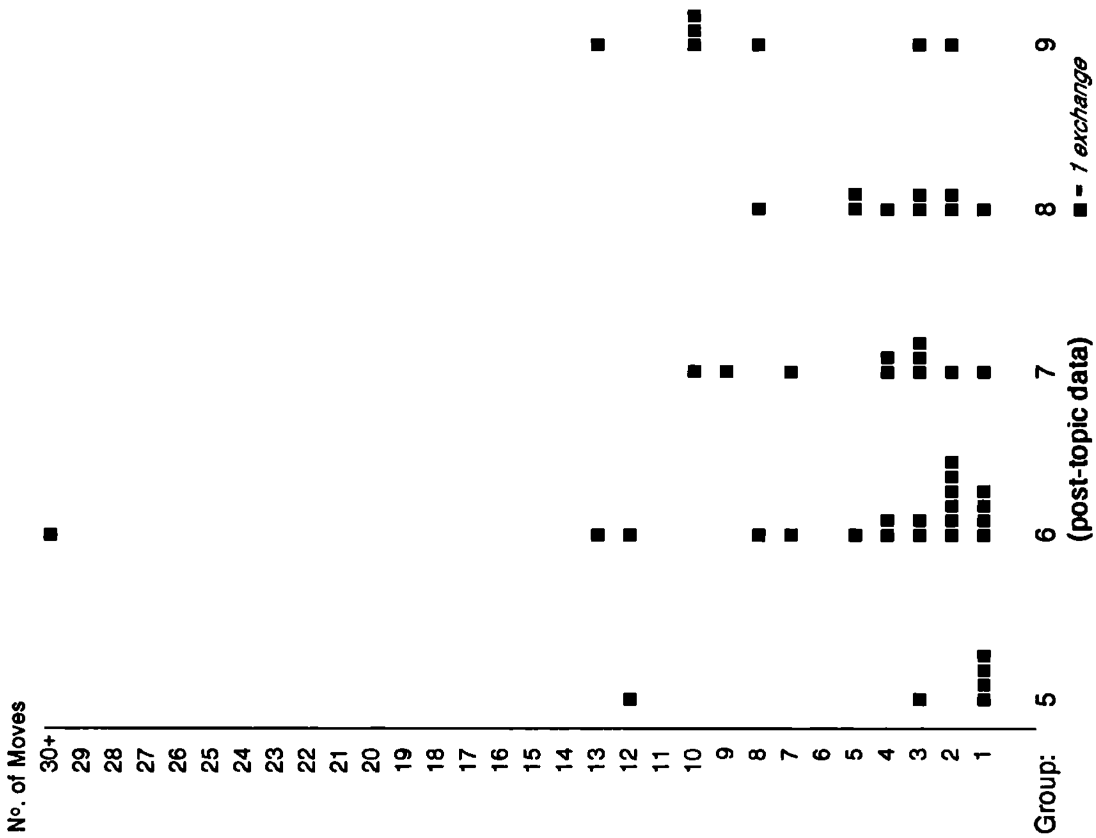
The figure shows the distribution of exchange length, and how this varied from group to group. Typically, there were several short exchanges of one or two moves, and in most cases there were also one or more quite lengthy exchanges. There is, though, no fully consistent pattern. For some groups, the distribution is very skewed towards the shorter exchanges. Others show a much flatter distribution. In most cases, though, the overall pattern for each group shows little variation between the pre-topic and the post-topic discussion. Hence this distribution is probably due as much to the particular mix of individuals in the group as to the characteristics of the activity. It is possible to relate the distribution of exchange length to the way the different groups worked together, thereby providing evidence to support the validity of this categorization of the data into exchanges.

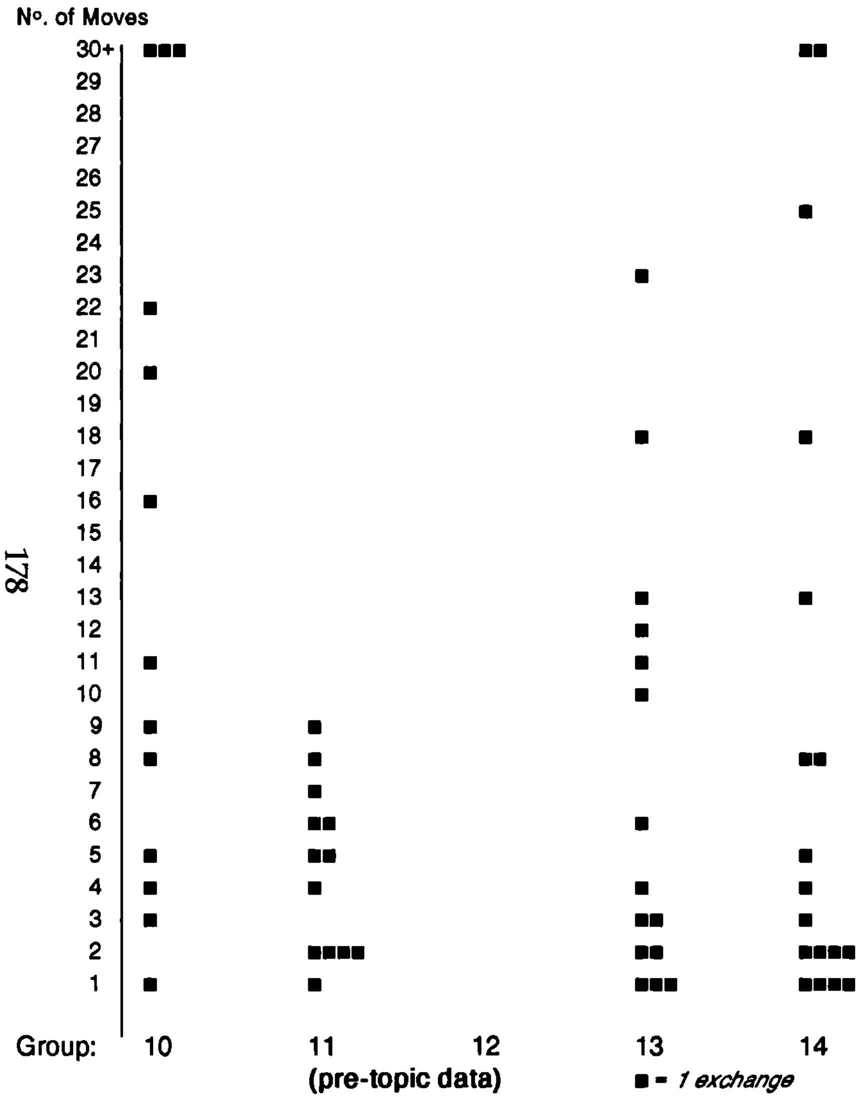
Group 4 (post-topic) was notable in that the audio tape indicated there was only limited task-oriented talk. Instead, the discourse for this group featured a high proportion of unproductive (or even counter-productive) chatter. This seemed to be due primarily to the influence of one charismatic individual and was possibly exacerbated by the presence of the tape recorder, which seemed to be intended as the direct recipient of much of this chatter. Consequently, it is unsurprising that the data reveal this group to have featured a preponderance of very short ideational exchanges. Group 5 seemed hardly to be discussing what they were going to incorporate in their map, either pre- or post-topic. They seemed instead simply to be working independently, drafting different versions of the map and then deciding whose to choose. It was around this that the majority of the discussion revolved, particularly for the post-topic map. Consequently, they engaged in very few ideational exchanges, and particularly in the post-topic session, these tended to be short. Group 8, like group 4, was prone to unproductive chatter, particularly in the post-topic session, when two additional group members were present. This change is reflected in the altered distribution of exchange lengths from pre- to post-topic. Thus holistic overviews of the way the groups worked together are reflected in this initial analysis of the exchange data.

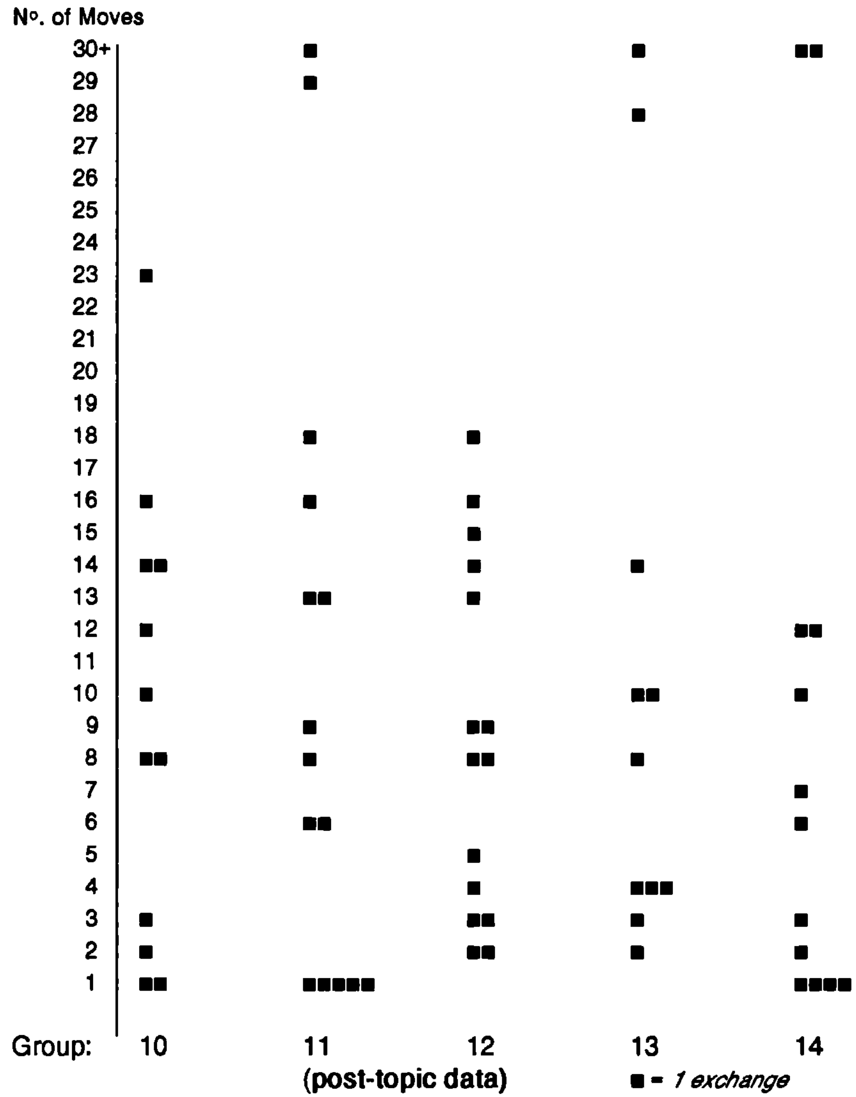
Figure 9.3: Number of Moves per Exchange




Nº. of Moves
30+ 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
Group: 5 6 7 8 9
(pre-topic data)
■ = 1 exchange


Nº. of Moves
30+ 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
Group: 5 6 7 8 9
(post-topic data)
■ = 1 exchange


Nº. of Moves
30+
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
Group: 10 11 12 13 14
(pre-topic data)
■ = 1 exchange


Nº. of Moves
30+
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
Group: 10 11 12 13 14
(post-topic data)
■ = 1 exchange

One preliminary hypothesis would be that the length of exchange is a measure of the amount of deliberation involved in deciding whether to include a construct or relationship in the map, and if so, the wording to be used to express it. Single move exchanges do not allow negotiation of meaning within the exchange. With these, whatever is initially stated must either be accepted or rejected as it stands, and any response would need to be tacit. Longer exchanges offer more possibilities. The members of the group can express explicit approval or rejection of the initial offering, or they can seek to modify it. Some examples will help to illustrate these possibilities.

Examples through the remainder of this chapter follow the format below, which is based on that used for the full transcript given in Appendix E, where full details are provided. The following columns are used here:

Actor Wording **Category** **Comment**

Overlapping wording is indicated by underlining and bracketing together, and here the beginnings of the overlapping utterances are aligned for clearer interpretation. The commentary supplies interpretive remarks where appropriate, drawing on cues provided by intonation and timing in the tapes.

Transcript 1: Group 11 (Sound and hearing; pre-topic)

P1	Yeah, this will go up here, I'm just drawing it, cos you know when you hear a guitar string, it vibrates, don't it ... when you leave it, right, so, guitar string, and then you've got, um	[Introducing]	
P2	<u>Yeah</u>	[Supporting]	

Here, P1 introduces an idea in what seems to be a fully determined form. P2's only contribution is to offer support with the single word "yeah". Following this, the group move on to a different topic and the connection is explored no further. There seems, therefore, to have been only a fleeting acknowledgement of the suggestion, and no real interest in it. No link is made on the concept map. Contrast this with the following exchange, in the same group.

Transcript 2: Group 11 (Sound and hearing; pre-topic)

P1	Right ... what do you think is ... what do you think's the important, one?	[Opening]	
P1	I think sound wave is, what do you think?	[Introducing]	
P2	Em ... yeah	[Supporting]	
P1	Sound wave	[Contextual]	Rehearsing the word
P1	What do you think, *** ?[19]	[Eliciting support]	
P3	[?]	[Supporting]	Actual words inaudible
P1	Yeah so we'll put sound wave at the top, yeah?	[Supporting]	

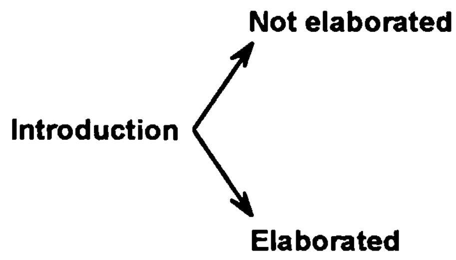
In this longer exchange, following the preliminary opening remark, P1 goes on to offer her idea of which construct is "most important". However, in doing so, she leaves it open for others to express support, and subsequently ensures that such support is affirmed, before reiterating the agreed outcome. This is then incorporated into the map. Thus, in this instance, there is a more self-conscious attempt to ensure agreement over the way the concept map is to develop, which in turn depends on shared understanding. Already it is emerging that the length of an exchange is only one factor in how productive that exchange is. The number of moves may, indeed, be only a correlate of underlying structural differences that determine whether scientifically acceptable meanings are adopted. Hence it is to an examination of the structure of ideational exchanges that we now turn.

9.4.2 Exchange Structure

Elaboration and non-elaboration

On examining the sequence of moves in the exchanges, it was possible to divide the exchanges into two distinct groups, based on how ideas were introduced and developed. The distinction may be represented thus :

19 *** = name deleted for anonymity (see Appendix E for a full explanation of the symbols)



The property of the exchange concerned here is whether the idea introduced is ***elaborated*** or ***not elaborated***. Elaboration was defined in Chapter 8 as a response that adds new propositional content to the discussion. Integrating moves are also elaborative, and were defined above as the linking or reconciliation of two sets of previously unrelated or conflicting ideas.

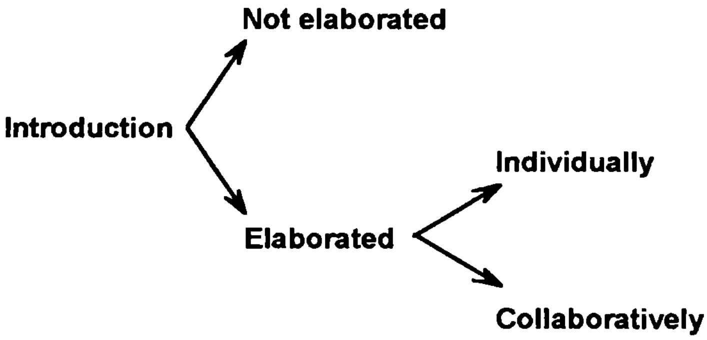
Not all exchanges featured elaborative moves. In some, an idea was introduced in essentially the same form in which it ultimately became either incorporated into the concept map, or alternatively ceased to be considered as a candidate for inclusion in the final map. Although in such cases there may have been a number of moves supporting the idea originally introduced, there was an absence of any moves which added to the information under consideration by the group.

In other cases, ideas were introduced in an incompletely developed form, and then, over the course of subsequent moves, further information relating to the original idea was brought into consideration. This phenomenon suggests that there may have been less of a grounding for the idea in shared understanding within the group, and that continued consideration of the idea required further information to be brought into play.

Individual and collaborative elaboration

Although whether an idea is elaborated upon during an exchange goes some way towards indicating whether the original idea was the subject of shared understanding, there is an important further consideration: did the children work together to establish the meanings incorporated in their map, or did they do this individually, with the others in the group acting more or less as spectators?

In some exchanges in the data, each step of elaboration was made by the same group member. If the others participated at all, it was generally by offering supporting moves. In other exchanges, the rest of the group were more active in generating the elaboration by extending the information under consideration. In the former cases, the elaboration may be said to have been achieved ***individually***, while in the latter cases it was achieved ***collaboratively***. Hence the schematic illustration of the development of an exchange gains a further bifurcation:



There are therefore three types of exchange structure, represented by the three routes (paradigms) through the diagram:

(i) Introduction → Not elaborated;

(ii) Introduction → Elaborated → Individually;

(iii) Introduction → Elaborated → Collaboratively.

To illustrate these contrasting possibilities, four examples will now be presented.

Transcript 3: Group 7 (Earth in space; pre topic)

P1	Oi, af, after star, shall we put s, the sun is a star?	[Introducing]	
P2	Yeah	[Supporting]	
P3	Under universe, we'll put, the sun is a star	[Contextual]	Spoken to tape recorder
P4	What's that?	[Query]	
P3	The sun is a star	[Answer]	
P2	Underneath star, we should put, sun, and you write, is a, is a thing	[Answer]	
P4	The sun is a star	[Contextual]	Rehearsing what has been said

Here, P1 proposes the next link in the concept map in a developed form. The following moves are concerned only with clarifying and approving the proposition prior to recording it, and the group then move on to the next topic of discussion. There is no questioning of the appropriateness of what P1 offers, and no attempt to elaborate upon it. This is simply a statement of what seems already to be well known by all in the group. This is one instantiation of the first paradigm, Introduction → Not elaborated.

An instantiation of the second paradigm, Introduction → Elaborated → Individually, follows.

Transcript 4: Group 10 (Sound; post topic)

P1	Sound wave travels from ear drum to cochlea	[Introducing]	
.			
.			
.			
P1	Sound [?] oh I forgot now	[Contextual]	Rehearsing what
P2	Sound wave	[Contextual]	to write on the
P1	Travels from your ear drum	[Contextual]	concept map
P3	To, to	[Contextual]	
P1	<u>Cochlea, then up your nervous system ...</u> up your nervous system to the brain	[Elaborating]	
P4	<u>Tra, sound waves travels, to, cochlea</u>	[Contextual]	

Here, the elaboration offered by P1 seems almost to be an afterthought, and is not pursued further, either through giving support or through further elaboration. Examples of this type of exchange are relatively rare in the data. On the other hand, there are many instantiations of the third paradigm, Introduction → Elaborated → Collaboratively. Two examples follow. The first is offered by way of contrast to the example in 6.3 above, in which the same link was discussed.

Transcript 5: Group 6 (Earth in space; post topic)

P1	Er, sun's a star?	[Introducing]	
P1	Sun's a star	[Supporting]	
P2	Sun's a star, plus a planet, so it could go there, isn't it?	[Elaborating]	
P2	It is a planet as well, isn't it? Isn't it a planet?	[Eliciting elaboration]	There seems to have been a silent Challenge
P2	Oh yeah you <u>can't go and visit the sun</u>	[Retracting]	
P1	<u>It's a star, isn't it,</u> and a star ain't a planet, so	[Elaborating]	
P2	So it's a star	[Supporting]	

Although this is only a brief exchange, it serves to bring P2's shaky understanding of the distinction between stars and planets into focus. In the subsequent moves, two kinds of additional information are brought into play. Firstly, P2 recalls that it is a property of stars that they cannot be visited, and a property of planets that they can. This is informal knowledge, only partially formed, and seems to be based on the idea that stars are fireballs, and consequently too hot for humans to approach. In fact, the group do not abandon this completely as a subject of conversation, and return to it later for a lengthy discussion on the nature of stars and planets (see 9.8). P1, on the other hand, employs a clear and distinct classification of celestial bodies in which stars and planets are exclusive categories.

The previous case, then, is an example of knowledge (that the sun is a planet) that had been taken to be shared, but that turned out to be problematic. In the following example, the members of the group seem actually to be striving to establish common understanding.

Transcript 6: Group 3 (Habitats; post topic)

P1	Food needs sunlight	[Introducing]	
P2	Hold on	[Contextual]	
P3	[?]	[Unclear]	Inaudible
P1	Well say it's a plant, then it needs sunlight	[Elaborating]	
P2	I suppose	[Hedging]	Sounds doubtful
P3	Food comes from plants	[Introducing]	
P1	Yeah, food comes from plants ... plants need sunlight	[Integrating]	
.			
.			
P2	Yeah but it doesn't all come from plants	[Challenging]	
P3	Yeah I know but, food does come from them ... food comes from other animals, doesn't it	[Elaborating]	Accepts objection but reaffirms. Tries to clarify
P4	Yeah ... yeah	[Supporting]	
P1	Other animals eat, plants	[Elaborating]	Takes up from P3
P3	Plants need sunlight	[Supporting]	

At the beginning of this series of moves, P2 appears not to understand the connection between sunlight and food, an important element in understanding energy flow in food chains. In subsequent moves each of the other participants builds on what has been said previously, with P2 assisting in the process by raising a further objection. The result is that there is achieved a complexity of relationships in the emerging concept map that does not seem to be due to any one participant alone.

The paradigms provide, then, a three-way, exhaustive classification of the ideational exchanges found in the data. Table 9.4 shows the breakdown of exchanges into the three types.

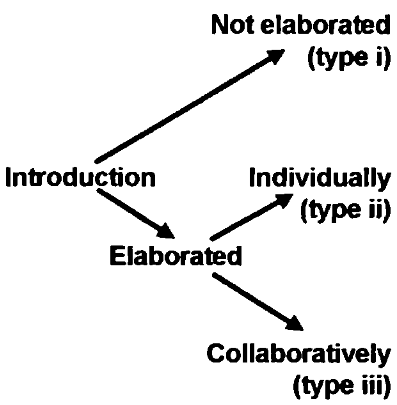
Table 9.4: Proportion of Ideational Exchanges of each Type

	Pre-topic				Post-topic			
Type of exchange ➡	(i)	(ii)	(iii)	Total	(i)	(ii)	(iii)	Total
Group ▼								
1					7 (35%)	3 (15%)	10 (50%)	20
2					13 (37%)	6 (17%)	16 (46%)	35
3	16 (53%)	4 (13%)	10 (33%)	20	16 (57%)	4 (14%)	8 (29%)	28
4	9 (60%)	3 (20%)	3 (20%)	15	11 (69%)	2 (13%)	3 (19%)	16
5	3 (43%)	1 (14%)	3 (43%)	7	5 (83%)	0	1 (17%)	6
6					12 (60%)	2 (10%)	6 (30%)	20
7	5 (42%)	0	7 (58%)	12	5 (56%)	1 (11%)	3 (33%)	9
8	7 (58%)	2 (17%)	3 (25%)	12	4 (50%)	0	4 (50%)	8
9	3 (27%)	1 (9%)	7 (64%)	11	4 (57%)	0	3 (43%)	7
10	3 (23%)	0	10 (77%)	13	1 (8%)	0	11 (92%)	12
11	6 (46%)	5 (39%)	2 (15%)	13	7 (47%)	1 (7%)	7 (47%)	15
12					6 (40%)	0	9 (60%)	15
13	5 (31%)	3 (19%)	8 (50%)	16	2 (18%)	2 (9%)	8 (73%)	11
14	7 (39%)	1 (6%)	10 (56%)	18	6 (46%)	2 (15%)	5 (38%)	13

Table shows, for each group, number and percentage of each exchange type within each concept mapping session.

In order to investigate whether the distribution of these exchanges represented stable features of the discourse, the relative proportions of the three types of exchange were compared for pre- and post-topic sessions across all the groups. Figure 9.4 shows the outcome from the crosstabulation.

Figure 9.4: Crosstabulation of Exchange Type by Session (All groups)



	Pre-topic	Post-topic	
Not elaborated (type i)	65 (44%)	92 (46%)	157 (45%)
Individually (type ii)	19 (13%)	21 (10%)	40 (12%)
Collaboratively (type iii)	63 (43%)	89 (41%)	152 (44%)
	147 (42%)	202 (58%)	349 (100%)

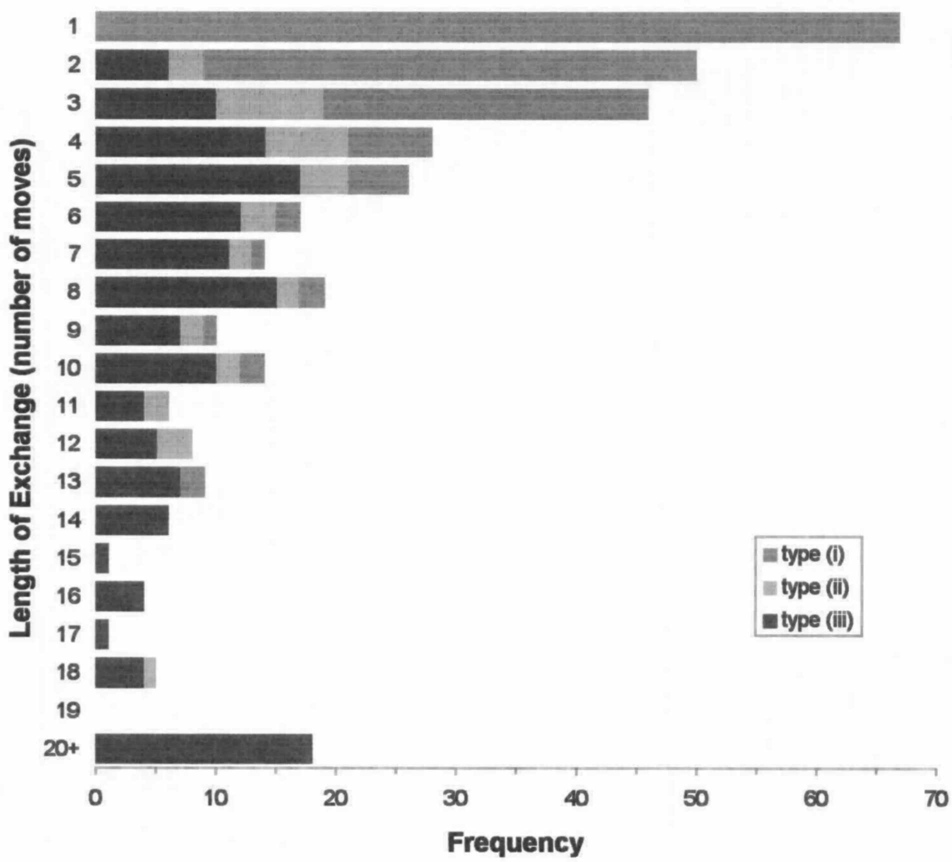
Chi-square: 0.54 (n/s)

Table shows number and row percentage of each exchange type with column and row totals

The crosstabulation and non-significant χ^2 show that the pattern of exchange types overall was consistent between pre- and post-topic sessions. This confirms the impression from the table that the individual groups varied only slightly in the distribution of exchange types between the two sessions.

Figure 9.5 displays the distribution of exchange length for the three different types of exchange. It reveals that the shorter exchanges were predominantly of type (i), with progressively greater proportions of type (iii) as the length increased. It does therefore appear that the length of an exchange is closely related to is structure.

Figure 9.5: Length of Exchange for Different Types



Based on data for 349 ideational exchanges

With these properties of ideational exchanges explained, attention can next be turned to which ones were associated with the establishment of scientific meanings within the groups.

9.4.3 Outcomes from Ideational Exchanges

The next step in the analysis was to link each exchange to its outcome, and hence to determine whether there was any relationship between the nature of this outcome and the properties of the exchange.

The object of engaging in an ideational exchange was to incorporate a construct term or a relationship between construct terms into the concept map. Evidence of what groups accepted as the outcome of an exchange is therefore provided by their concept maps. The decision to add a term or link to the concept map was a decision to "go public", as the map could be read by the teacher and the researcher, and, in the case of some of the groups, discussed in front of the class. Each exchange could result in either the incorporation into the map of some formulation of the idea under consideration, or with the rejection of that idea, in which case the relationship discussed would be absent from the concept map.

As described in Chapter 8, the links made in the final concept map were coded in terms of the kind of knowledge portrayed. The possible outcomes from an ideational exchange were:

- a scientifically acceptable relationship;
- a vague or inexplicit connection, or a usage acceptable in everyday talk, but having no scientifically acceptable meaning;
- a relationship that was scientifically incorrect, and represented no acceptable everyday usage (that is, a misconception);
- no connection at all.

Since, in the course of each exchange, the children referred to the terms in the concept map, it was possible to relate each exchange directly to a corresponding feature of the finished map. With these analyses completed, a computer data file was set up to record, for each exchange, its properties, its outcome in the concept map and information about the group, topic and session. This enabled statistical analysis of any desired combination of these variables, using SPSS PC+ software (Norusis, 1988).

It may be expected that the longer an exchange, the more opportunity there is for an agreed outcome to be established, and given that the group members may have differing viewpoints, the more likely it is that the shared conclusion will be scientifically acceptable. The data revealed the following descriptive statistics (Table 9.5):

Table 9.5: Number of Moves for Exchanges by Outcome

Outcome	Number of cases	Mean length (in moves)	Standard deviation	95% Conf. interval
Scientific	190	8.31	9.85	±1.40
Other	159	4.54	5.89	±0.92

Based on data for 349 ideational exchanges

It appears, from these data, that the mean length of those exchanges that resulted in a scientifically acceptable outcome did indeed exceed the mean length of less successful exchanges by a substantial amount. But it may also be that the type of exchange affects the outcome. Figures 9.6 to 9.9 portray the relationship between the type of exchange and its outcome in the concept map.

Figure 9.6: Outcomes from Ideational Exchanges (Topic 1)

Topic 1: Habitats	Outcome in Concept Map			
	Scientific meaning	Everyday meaning	Misconception	Not adopted
Introduction → Not elaborated (type i)	*PRE:* 6 *POST:* 18	*PRE:* - *POST:* 3	*PRE:* 1 *POST:* -	*PRE:* 18 *POST:* 26
Introduction → Elaborated → Individually (type ii)	*PRE:* 3 *POST:* 12	*PRE:* - *POST:* -	*PRE:* 1 *POST:* -	*PRE:* 3 *POST:* 3
Introduction → Elaborated → Collaboratively (type iii)	*PRE:* 10 *POST:* 32	*PRE:* - *POST:* 2	*PRE:* - *POST:* -	*PRE:* 3 *POST:* 3

Figure 9.7: Outcomes from Ideational Exchanges (Topic 2)

Topic 2: Earth in space	Outcome in Concept Map			
	Scientific meaning	Everyday meaning	Misconception	Not adopted
Introduction → Not elaborated (type i)	*PRE:* 6 *POST:* 18	*PRE:* - *POST:* 2	*PRE:* 3 *POST:* 2	*PRE:* 9 *POST:* 8
Introduction → Elaborated → Individually (type ii)	*PRE:* 2 *POST:* 2	*PRE:* 2 *POST:* -	*PRE:* - *POST:* -	*PRE:* - *POST:* 1
Introduction → Elaborated → Collaboratively (type iii)	*PRE:* 14 *POST:* 13	*PRE:* 2 *POST:* -	*PRE:* - *POST:* -	*PRE:* 4 *POST:* 4

Figure 9.8: Outcomes from Ideational Exchanges (Topic 3)

Outcome in Concept Map

Topic 3: Sound and hearing

Introduction

Elaborated

	Scientific meaning	**Everyday meaning**	**Misconception**	**Not adopted**
Not elaborated (type i)	*PRE:* 3 *POST:* 4	*PRE:* 1 *POST:* 1	*PRE:* - *POST:* -	*PRE:* 18 *POST:* 10
Individually (type ii)	*PRE:* 3 *POST:* 2	*PRE:* 1 *POST:* -	*PRE:* - *POST:* -	*PRE:* 4 *POST:* 1
Collaboratively (type iii)	*PRE:* 18 *POST:* 24	*PRE:* - *POST:* 2	*PRE:* 1 *POST:* 1	*PRE:* 11 *POST:* 8

Figure 9.9: Outcomes from Ideational Exchanges (All Topics)

Outcome in Concept Map

All Topics

Introduction

Elaborated

	Scientific meaning	**Everyday meaning**	**Misconception**	**Not adopted**
Not elaborated (type i)	*PRE:* 15 *POST:* 40	*PRE:* 1 *POST:* 6	*PRE:* 4 *POST:* 2	*PRE:* 45 *POST:* 44
Individually (type ii)	*PRE:* 8 *POST:* 16	*PRE:* 3 *POST:* -	*PRE:* 1 *POST:* -	*PRE:* 7 *POST:* 5
Collaboratively (type iii)	*PRE:* 42 *POST:* 69	*PRE:* 2 *POST:* 4	*PRE:* 1 *POST:* 1	*PRE:* 18 *POST:* 15

Figures based on data for 349 ideational exchanges

The distribution of the different types of move and of outcome varied over the three topics, but there were some trends that may be noted. There were fairly large numbers of exchanges in which the idea introduced did not find its way into the concept map, as shown by the right hand column. Inasmuch as a pattern may be discerned, these tended to be type (i) exchanges. Where there was a scientific meaning adopted, the tendency was for this to be the outcome from a type (iii) exchange, although in topic 2's post-topic session there was a sizeable proportion of scientific outcomes from type (i) as well.

To make a first step in understanding how interaction contributes to the construction of scientific meanings, it was necessary to investigate whether it is the *amount* or the *type* of discussion that is more influential in establishing scientifically appropriate meanings in the group's concept map. For this purpose, stepwise multiple regression was used to identify the independent variable that best "explained" the variance in the dependent variable. A "dummy coded" independent variable (Kerlinger & Pedhazur, 1973) was set up to represent the type of exchange (type ii exchanges were excluded as they were too few in number). The length of the exchange (in moves) was the second independent variable. The dependent variable represented whether or not a scientifically appropriate relationship was incorporated in the concept map as a result of the exchange, and was dichotomous.

The regression programme entered exchange type[20] but not exchange length[21] as predicting the outcome from an exchange. It seems, therefore, that it is the nature of the exchange that determines whether it is likely to result in a scientifically acceptable outcome, rather than the length of the exchange. Collaboratively elaborated exchanges were, over the whole data set, more likely to give a positive result. Here, then, is further evidence that collaborative establishment of meaning has a positive influence on the content of a concept map.

On the strength of these findings, a more detailed analysis of the interaction between exchange type and outcome was warranted. The purpose of this was to identify where significant effects were occurring, using χ^2 and

---

[20] $p < 0.0005$

[21] $p = 0.32$

related indices of association. For the purposes of this next stage in the analysis, it was necessary to collapse some of the categories from Figures 9.6 to 9.9 together, as expected frequencies for some cells were low Accordingly, the concept map outcome was reclassified as "scientific" versus "non-scientific" (comprising everyday meaning, misconception or not adopted). For crosstabulations involving individual topics, the exchange type also had to be collapsed into "individualistic" (type i and ii) or "collaborative" (type iii). Where expected frequencies fell between ten and five, Yates' correction for continuity was applied (Guilford, 1981). Figure 9.10 shows the results of this crosstabulation for all the topics together.

It is evident that the pattern identified in the cell frequencies was unlikely to be random. A substantial component of the deviation from the expected frequencies (as indicated by the standardized residuals shown in the cells) is located along the axis from top left to lower right of the contingency table. The relationship was not particularly strong (as shown by the value of Cramér's *V*) but was sufficient to suggest that the relationship had practical significance. The tendencies identified in the preceding discussion were due to more than chance.

Turning next to each topic separately, Figure 9.11 shows the crosstabulations for these. The ϕ coefficient is helpful here for comparing effects across the different topics, as, unlike χ^2, it is not affected by sample size. Using this as a guide, some interesting patterns emerge. Firstly, the relationship between exchange type and outcome is noticeably stronger for pre- than for post-topic sessions. Although the results for topic 2 do not attain statistical significance, they nevertheless follow this pattern. One implication from this would be that the state of children's knowledge about the topic and familiarity with the terms to be mapped were factors in the relationship. Increased confidence in their knowledge meant that the children needed less negotiation to achieve an appropriate outcome. As a consequence, the individualistic/scientific cell is larger for topics 1 and 2, post-topic. This is not so for topic 3, but then it must be recalled that the teacher changed some of the terms in the concept map for the second session in this topic.

There seem, therefore, to be initial indications that the level of "demand" associated with terms presented to the children affects the extent to which collaborative elaboration is needed to settle on the links to make. Nevertheless, we should not undervalue the post-topic discussion. The substudy has demonstrated that collaboration still has an effect on the post topic map, and for two of the three topics, there was still a clear relationship between concept map outcomes and whether those outcomes were developed through collaboration.

Summary and conclusions

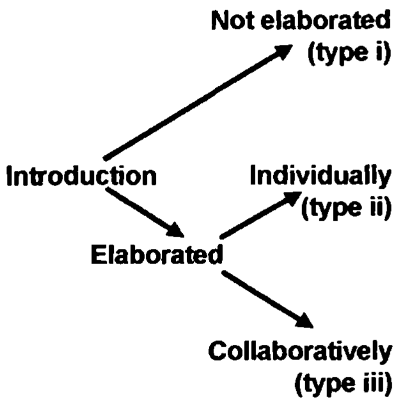
Three distinctive exchange structures have been identified that characterize the patterns of discourse within the groups. One feature of these exchanges seems to be significant in promoting the use of scientific language: when children share the task of elaborating the meaning of a relationship in the concept map, they are more likely to produce scientifically relevant links in their concept map. In such cases, the meaning that is finally accepted may take shape over several turns in the discussion. Evidence is beginning to accumulate that collaborative concept mapping facilitates a distinctive and, moreover, productive form of discussion.

There are also indications that the kind of terms the children are working with, and the level of children's understanding of those terms, are factors that influence the extent to which collaboration is needed to agree on suitable links between the terms.

These findings motivate a closer examination of the rôle of the different types of exchange in creating the map, and the ways in which they interact with the raw materials for the map: the set of construct terms.

Figure 9.10: Crosstabulations of Outcome by Exchange Type

a) All topics; pre-topic session

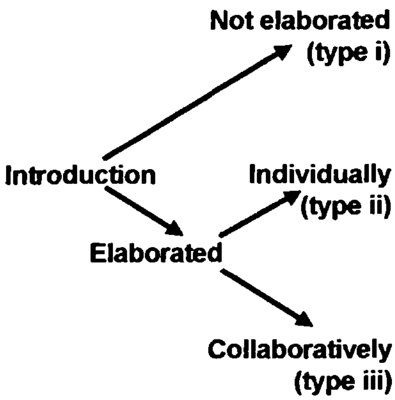


	Non-scientific outcome	Scientific outcome	
Not elaborated (type i)	50 (76.9%) *2.3*	15 (23.1%) *-2.6*	65 (44.2%)
Individually (type ii)	11 (57.9%) *0.1*	8 (42.1%) *-0.1*	19 (12.9%)
Collaboratively (type iii)	21 (33.3%) *-2.4*	42 (66.7%) *2.7*	63 (42.9%)
	82 (55.8%)	65 (44.2%)	147 (100%)

Chi-square: 24.68 ($p < 0.0005$)

Cramérs V: 0.41

b) Post-topic session



	Non-scientific outcome	Scientific outcome	
Not elaborated (type i)	52 (56.5%) *2.9*	40 (43.5%) -2.2	92 (45.5%)
Individually (type ii)	5 (23.8%) *-1.1*	16 (76.2%) *0.8*	21 (10.4%)
Collaboratively (type iii)	20 (22.5%) *-2.4*	69 (77.5%) *1.9*	89 (44.1%)
	77 (38.1%)	125 (61.9%)	202 (100%)

Chi-square: 24.27 ($p < 0.0005$)

Cramérs V: 0.35

Tables show number row per cent, standardized residual,
with column and row totals
Based on data for 349 ideational exchanges

Figure 9.11: Crosstabulations of Outcome by Exchange Type (Individual Topics)

a) Topic 1; pre-topic

	Non-scientific outcome	Scientific outcome	
Individualistic (type i, type ii)	23 (71.9%) *1.0*	9 (28.1%) *-1.2*	32 (71.1%)
Collaborative (type iii)	3 (23.1%) *-1.6*	10 (76.9%) *1.9*	13 (28.9%)
	26 (57.8%)	19 (42.2%)	45 (100%)

Chi-square: 7.13 (p = 0.0076))

Phi: 0.45

b) Topic 1; post-topic

	Non-scientific outcome	Scientific outcome	
Individualistic (type i, type ii)	32 (51.6%) *1.8*	30 (48.4%) *-1.4*	62 (62.6%)
Collaborative (type iii)	5 (13.5%) *-2.4*	32 (86.5%) *1.8*	37 (37.4%)
	37 (37.4%)	62 (62.6%)	99 (100%)

Chi-square: 14.37 (p = 0.0002)

Phi: 0.38

c) Topic 1; overall

	Non-scientific outcome	Scientific outcome	
Individualistic (type i, type ii)	55 (58.5%) *2.2*	39 (41.5%) *-1.9*	94 (65.3%)
Collaborative (type iii)	8 (16.0%) *-3.0*	42 (84.0%) *2.6*	50 (34.7%)
	63 (43.8%)	81 (56.3%)	144 (100%)

Chi-square: 23.97 (p < 0.0005)

Phi: 0.41

Tables show number row per cent, standardized residual with column and row totals *d) Topic 2; pre-topic*

	Non-scientific outcome	Scientific outcome	
Individualistic (type i, type ii)	14 (63.6%) *1.1*	8 (36.4%) *-1.0*	22 (52.4%)
Collaborative (type iii)	6 (30.0%) *-1.1*	14 (70.0%) *1.1*	13 (28.9%)
	20 (47.6%)	20 (47.6%)	42 (100%)

Chi-square: 3.50 (n/s)

Phi: 0.34

e) Topic 2; post-topic

	Non-scientific outcome	Scientific outcome	
Individualistic (type i, type ii)	13 (39.4%) *0.5*	20 (60.6%) *-0.4*	33 (66.0%)
Collaborative (type iii)	4 (23.5%) *-0.7*	13 (76.5%) *0.5*	17 (34.0%)
	17 (34.0%)	33 (66.0%)	50 (100%)

Chi-square: 0.65 (n/s)

Phi: 0.16

f) Topic 2; overall

	Non-scientific outcome	Scientific outcome	
Individualistic (type i, type ii)	27 (49.1%) *1.0*	28 (50.9%) *-0.9*	55 (59.8%)
Collaborative (type iii)	10 (27.0%) *-1.3*	27 (73.0%) *1.0*	37 (40.2%)
	37 (40.2%)	55 (59.8%)	92 (100%)

Chi-square: 4.48 (p = 0.0343)

Phi: 0.22

Tables show number row per cent, standardized residual with column and row totals *g) Topic 3; pre-topic*

	Non-scientific outcome	Scientific outcome	
Individualistic (type i, type ii)	24 (80.0%) *1.4*	6 (20.0%) *-1.7*	30 (50.0%)
Collaborative (type iii)	12 (40.0%) *-1.4*	18 (60.0%) *1.7*	30 (5%)
	36 (60.0%)	24 (40.0%)	60 (100%)

Chi-square: 10.00 (p = 0.0016)

Phi: 0.41

h) Topic 3; post-topic

	Non-scientific outcome	Scientific outcome	
Individualistic (type i, type ii)	12 (66.7%) *1.5*	6 (33.3%) *-1.3*	18 (34.0%)
Collaborative (type iii)	11 (31.4%) *-1.1*	24 (68.6%) *0.9*	35 (66.0%)
	23 (43.4%)	30 (56.6%)	53 (100%)

Chi-square: 4.66 (p = 0.0309)

Phi: 0.34

i) Topic 3; overall

	Non-scientific outcome	Scientific outcome	
Individualistic (type i, type ii)	36 (75.0%) *2.2*	12 (25.0%) *-2.3*	48 (42.5%)
Collaborative (type iii)	23 (35.4%) *-1.9*	42 (64.6%) *2.0*	65 (57.5%)
	59 (52.2%)	54 (47.8%)	113 (100%)

Chi-square: 17.37 (p < 0.0005)

Phi: 0.39

Tables show number, row per cent, standardized residual with column and row totals

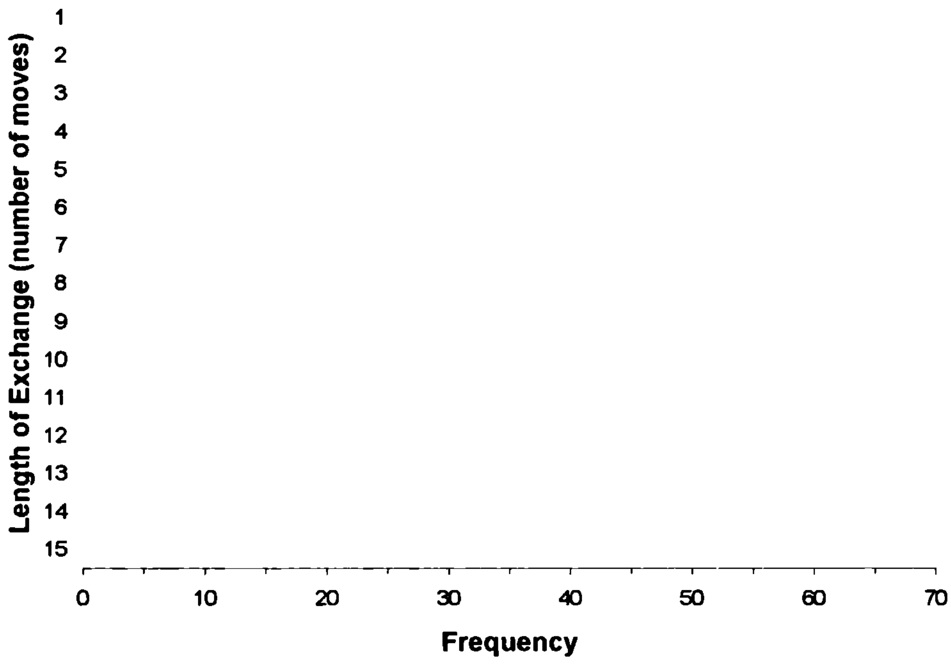
9.5 The Functioning of Ideational Exchanges

In this part of the chapter, a more interpretive analysis is made of how the different types of exchange functioned in the overall process of creating a group concept map. This involves examining examples of each type of exchange in turn, and drawing out their rôle in the task as a whole. The underlying purpose is to identify ways that the task is supportive of learning in science, and if possible, ways that it may be made more so.

9.5.1 Type (i) Exchanges

Exchanges of type (i) (introduction → not elaborated) were relatively common in the data. They could range in size from a single introductory move up to ten moves or more. Most were fairly short, consisting of about two or three moves. Figure 9.12 portrays these data.

Figure 9.12: The Length of Type (i) Exchanges



Mean length: 2.39 Median: 2.00 Std Dev: 2.11

Based on data for 157 exchanges

Examining how these short, non-elaborated exchanges functioned revealed the first important features of the discourse.

The starting point for determining the function of these exchanges was a pair of contrasting possibilities. On the one hand, it could be that in type (i) exchanges individual pupils were introducing candidate relationships in a sufficiently developed form for them to be included in the concept map without modification. Any following moves would therefore be either supportive, prior to adding the relationship to the map, or possibly challenging, prior to rejecting the suggestion. The alternative possibility is that type (i) exchanges were opening bids: candidate preliminary moves to introduce an idea for discussion, but for which further elaboration did not take place for some reason. In such cases, the relationship initially introduced would be less explicit. Thus the form in which an idea was introduced in an exchange is one key to determining what the function was of the exchange within the discourse as a whole.

Different ways of introducing an idea vary in their degree of determinacy. The idea could be in the form of a single construct term, a pair of terms, or two or more terms connected by some sort of implicit or explicit relationship. These possibilities may be illustrated by the following paradigm examples, which are placed in the context of a concept map for topic 1, Habitats.

Exploratory talk

- "Animal next"
- "Animal and predator"
- "Animal connects to predator"
- "Animals are predators "

Final draft

These forms show a gradation in explicitness, with a greater tendency towards what Barnes (1976) calls "final draft" in the last example, as indicated by the scale on the right. These different degrees of determinacy would be expected to govern the scope for appropriating the idea in different ways and of elaborating upon it further.

There were many instances in the data that seemed to indicate the pupils were indeed making fully formed, "final draft" statements that could have been transferred directly onto the map without further discussion. The following short exchange followed on from a discussion about the relationship between sound wave and ear drum. The children had just concluded that there was a link "because you hear sound waves".

Transcript 7: Group 12 (Sound and hearing; post-topic)

P1	And guitar string makes sound waves	[Introducing]
P2	Sound waves	[Supporting]

There is no further exploration of this ready-formed idea, only agreement. In the following example, there is *only* the single introductory move before the relationship is written onto the concept map.

Transcript 8: Group 4 (Habitats; post-topic)

P Predators have competition over food

Type (i) exchanges wherein the nature of the proposed relationship was explicit were in the majority. For both pre- and post-topic sessions, they outnumbered those in which the relationship was unclear by about two to one. However, it would not be true to conclude that these were always cases of dogmatic assertions. Sometimes, the relationship was explicitly stated in the introducing move, but was said in such a tone as to give it the effect of a question inviting support, as in the next example.

Transcript 9: Group 9 (Earth in space; post-topic)

P Next we can write moon, moon goes round the Earth?

Thus what was offered was marked as provisional. Hence it was not only the explicitness in the relationship expressed that determined what happened subsequently. The distinction between exploratory talk and assertion did not coincide with how the introducing move was formulated because control over the discussion did not rest with the person introducing the idea. Even if one of the children was presenting what was presumably intended to be a definitive proposition, its status could be, and frequently was, redefined as others in the group reacted in differing ways: ignoring it; hedging; acknowledging it but moving on. The next extract shows this process in action.

Transcript 10: Group 4 (Habitats; pre-topic)

Te	Right, so if you put ... plants ... and put a circle round it ... then sunlight	[Contextual]	Explaining how to write on the map
P1	Plants ... need ... sunlight	[Contextual]	Rehearsing while writing
P2	They grow out the ground	[Introducing]	
P3	I've just thought of something ... plant if it's plants then the plants are food to animals, the animals need the food for energy, and survival	[Introducing]	

Here, P2's (rather feeble) suggestion is bypassed in favour of the more developed and more relevant set of relationships proposed by P3. P3's proposal is progressive; it opens up further possibilities and consequently is taken forward. P2's proposal, on the other hand, does not have such potential, and consequently degenerates. There are faint parallels here with the Lakatosian notion of research programmes (Lakatos, 1970): a new idea is taken up, not so much because it is right while its predecessor is wrong, but because it offers more scope for development.

As a consequence of this group effect, many of the ideas raised in these exchanges were never adopted, and were only fleetingly supported by the proponent or the others in the group. There were clear indications in some cases that offerings were to be regarded as tentative, as a kind of thinking aloud. In the following example, a relationship is introduced and then quickly retracted by the same group member.

Transcript 11: Group 1 (Habitats; post-topic)

P	*Survival is needed against predat-*	[Introducing]
P	No	[Retracting]

All of the preceding examples begin with an explicit first move. In contrast to the above cases, there were other exchanges in which a member of the group suggested a connection between two construct terms in a vague way, apparently offering this to the group as a possibility for inclusion in the map. By leaving the nature of any connection unstated, it was possible to introduce the idea as it occurred to the person, without the obligation to think it through further at that stage. This kind of introduction may therefore be read as "I think there is a connection here. Shall we pursue it?". In some cases, the inquiring nature of the introducing move was quite plain, as here:

Transcript 12: Group 14 (Sound and hearing; post-topic)

P	You know these sounds ... do you think we should take, do you think we should link them up on the ear drum or the, cochlea?

With a potential link brought into the public domain, it becomes possible for the resources of the group as a whole to be brought to bear in evaluating it. When there was no direct response to the idea in the form of supporting moves, this was not an indication that the offering was pointless or unwelcome. As with the more explicitly introduced links, frequently the reason that a suggestion was not taken up was that another group member, scanning the list of available construct terms, offered a more attractive suggestion shortly afterwards. Such was the case in the example below.

Transcript 13: Group 2 (Habitats; post-topic)

P1	Survival, food	[Introducing]	
P2	Energy comes from plants	[Introducing]	Surprised intonation, suggesting this is seen as an insight
P1	Hmmm	[Hedging]	
P3	Plants	[Supporting]	
P4	Plants	[Supporting]	
P1	Energy	[Supporting]	
P1	OK	[Supporting]	

Here, the idea offered by P2 is more specific than P1's vague association, which is not explored further and so remains a single move exchange. The others use hedging and supporting moves to retain the second idea offered as the subject of the discourse while they consider its usefulness, and then go on to accept it. (In fact, they later go on to query whether it is appropriate to predicate "energy comes from" of plants, as they recognize the sun as providing energy.) The second of the two exchanges featured in this excerpt shows how longer type (i) exchanges could serve as a way of exploring possibilities.

At some points in the data, there featured a cluster of short, non-elaborative exchanges in close succession, as the children proposed a range of possibilities for the next link in the concept map. The following example illustrates this.

Transcript 14: Group 2 (Habitats; post-topic)

P1	Light is needed for survival	[Introducing]	
P2	Hmmm	[Hedging]	
P3	Hmmm ... no	[Retracting]	
P1	Humans ... humans	[Introducing]	
P2	Humans [?]	[Contextual]	Probably a joke
P4	No ... em ... light ... is prod ... sunlight is produced in the day ... during the day	[Introducing]	
P?	Sunlight	[Supporting]	
P1	What word would you start from?	[Eliciting elaboration]	Unsure how this fits
P3	Day	[Supporting]	
P4	Day	[Supporting]	
P1	Sunlight grows plants	[Introducing]	
P4	Sunlight is produced ... yeah sunlight grows, helps grow plants	[Elaborating]	

This group are working out possible ideas around the term "sunlight". The earliest offerings do not meet with much support. It is P1's "sunlight grows plants" that attracts the attention of the group, and P4 switches from supporting his own previous suggestion to elaborate on this. This episode develops through having to hand the set of construct terms, from which the children are trying to make appropriate links to what is already on the map. They look for potential links between pairs of terms, and introduce these for further consideration.

The phenomenon of multiple type (i) exchanges in succession was evident in some of the discussions for topic 1 (Habitats) and for topic 3 (Sound and hearing), but not for topic 2 (Earth in space). It was therefore worth looking more closely for an explanation for this difference. The construct terms chosen by the teacher for topic 2 were distinctive in that, with the exception of "universe", they were categories of object in space. Hence there was only a relatively limited number of different ways in which these terms could be related (primarily, class inclusion and spatial relationships). The terms chosen for the other topics exhibited more variability. Along with classes of observable object, there were also phenomena that were not directly observable ("sound wave") and more abstract theoretical constructs ("competition", "survival"). These afforded a wider range of relationships, such as cause and effect, part-whole, characteristics and dependencies. These latter two topics would therefore be expected to generate more discussion about potential relationships than topic 2, in which the scope for combining constructs in different ways was more limited.

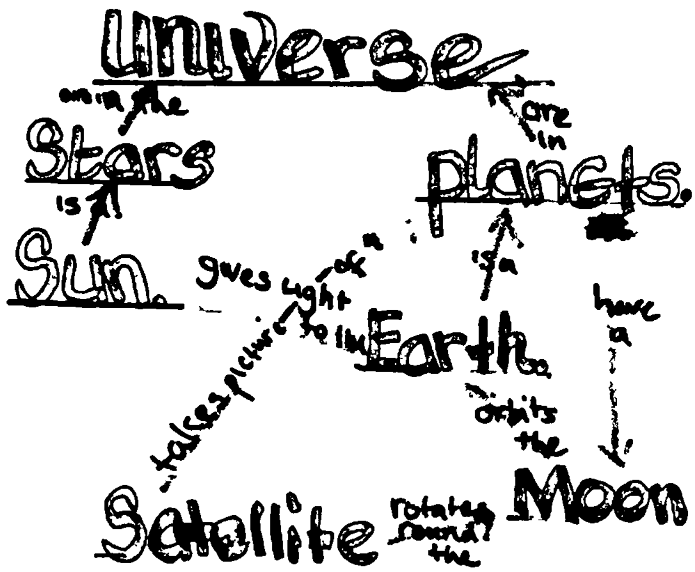
It is important, therefore, to view each individual exchange as functioning within the wider context of the activity as a whole. From this viewpoint, type (i) exchanges could perform the potentially valuable function of opening up a range of possibilities for further consideration. The children were trying out ideas in an environment where the power distribution allowed tentative, ill-formed ideas to be voiced, and where even well-formed ideas were negotiable. Short, non-elaborative exchanges were symptomatic of the breadth of possible connections that the children explored.

These exchanges had an outcome. The outcome was either a link in the children's concept map, or the absence of such a link, if the children chose not to pursue the idea. Commensurate with this hypothesised exploratory function, the majority of type (i) exchanges did not lead to a link being made (see Figure 9.9). This was true across the topics overall. Within the individual topic areas, topic 2 (Earth and space) did not follow this pattern exactly. For topic 2's post-topic mapping session, most of the type (i) exchanges resulted in scientifically acceptable connections in the map, whereas the pattern found in all other topic-sessions, was for the largest cell entry to be "not adopted". However, it was still the case that most of the ideas that were not adopted resulted from type (i) exchanges, for all topics and sessions (Figure 9.9).

An explanation for why the discussion for topic 2 did not conform to the overall pattern can be found by comparing the pre- and post-topic maps for the groups. For two groups, 7 and 9, there was very little difference between the pre-topic and post-topic maps, suggesting that there was nothing new to discuss in the second session. (Compare, for example, group

7's post-topic concept map shown in Figure 9.13 with the pre-topic map shown in Appendix E).

Figure 9.13: Post-topic Concept Map (Group 7)



For both these groups, the post-topic discussion featured an unusually high proportion of type (i) exchanges that resulted in scientifically correct links in the maps. This would suggest that often the children in groups 7 and 9 were simply repeating links they had established in the previous session, without feeling any need to negotiate these. Other groups also showed few signs that, in the post-topic session, there was much left to explore in the way of which term related to which. It seems that this was a case in which the level of demand provided by the concept maps was not sufficiently high to provoke the more productive forms of discourse that characterized the other sessions recorded. This was no doubt related to the characteristics of the construct terms chosen for the topic. As was observed above, the number of different ways the terms for this topic could be combined were limited in comparison to the other topics featured in this research.

Thus type (i) exchanges were not all symptomatic of socially constructed exploration. There were also cases in which there was no exploratory talk. In some of these cases, acceptable ideas were suggested, and adopted without further ado. In others, promising leads were not pursued, or unscientific language went unchallenged. P4's proposition "sunlight is produced during the day" in Transcript 14 above is an example of a statement of "everyday" understanding that does not get explored further, and which ultimately is revisited and incorporated into the concept map. Such a link does nothing to clarify the scientific relationships in the map. In the next example, the children fail to take up an idea that could, if taken on further, have opened up the discussion to a key idea in the understanding of ecosystems.

Transcript 15: Group 4 (Habitats; pre-topic)

P1	Energy is part, energy is part of the food chain	[Introducing]
P2	What have you done, you've gone ahead of us	[Contextual]
P3	Yeah, you've gone ahead of us	[Contextual]

The others in this group have not kept up with the flow of the discourse, and it was probably the case that this particular group was not working in such close cooperation as it might have done (see the discussion in 9.3.3 above). The example below shows how a misconception can be introduced into the discussion, and ultimately can be written onto the map unchallenged.

Transcript 16: Group 6 (Earth in space; post-topic)

P1	Moon ... moon's a planet, isn't it?	[Introducing]	
P2	What?	[Query]	
P3	What's been happening in your class?	[Contextual]	P3 has been absent
P1	Moon's a planet	[Supporting]	

Such cases were rare, but cannot be ignored. Figure 9.9 shows that the majority of misconceptions in the concept maps, as well as a large proportion of the everyday meanings, arose out of type (i) exchanges. We shall return to the subject of misconceptions in 9.6. Meanwhile, we must note that group concept mapping is by no means a panacea for difficulties in learning science. In cases like the latter one, there is a very public outcome in terms of an invalid link in the concept map that can alert the teacher to the problem. Missed opportunities are harder to detect, although there remains the possibility of the teacher's suggesting a connection that could be followed up where the pupils have not done so.

Summary and conclusions

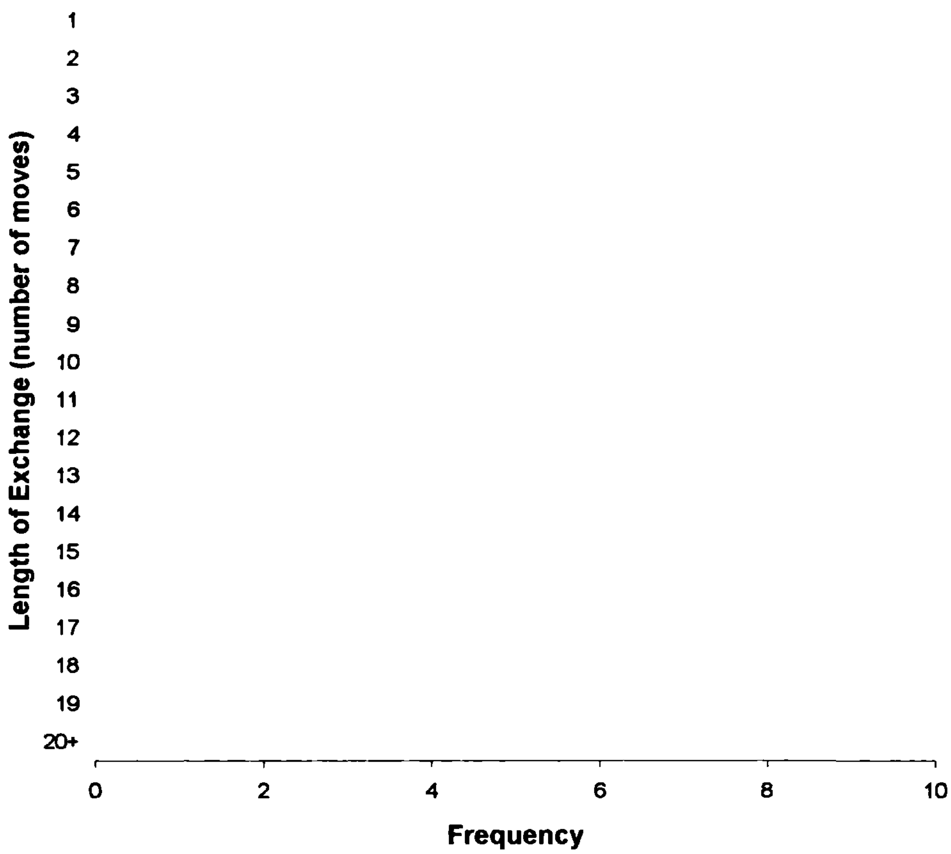
The existence of type (i) exchanges in the data is a sign that the children were introducing tentative candidates for inclusion in the concept maps. Whether or not an idea was taken as tentative was, however, socially determined, and not the decision of one group member. Sometimes, the tentativeness of an idea was imposed by the originator, either explicitly, or through expressing the nature of the linking relationship in a vague way. At other times, the group redefined the status of an idea that was originally introduced without any appearance of its being provisional.

On occasions, however, type (i) exchanges could be indicative of ready-formed ideas' being introduced into the map. The extent to which the children were encouraged to explore provisional links was apparently influenced by the nature of the construct terms they were given, and how these were related in the domain in question. Exploratory talk was associated with sets of terms that were ontologically dissimilar, and consequently related to each other in varying ways. Ready-formed ideas were more likely in cases where the terms were of similar kinds, and related to each other in a restricted range of ways.

9.5.2 Type (ii) Exchanges

As shown above, examples of type (ii) exchanges (introduction → elaborated → individually) were relatively rare. In such an exchange, the idea is developed primarily by the same individual who introduces it. Others in the group may make a contribution, by asking for elaboration, or by expressing support for the developing idea. Type (ii) exchanges tended to be of short to medium length, as displayed in Figure 9.14.

Figure 9.14: The Length of Type (ii) Exchanges



Mean length: 5.95 Median: 5.00 Std Dev: 3.70

Based on data for 40 exchanges

In the following example of a type (ii) exchange, the group begin by discussing trees' production of oxygen and animals' need of oxygen. This seems to remind one of the children about the process of photosynthesis.

Transcript 17: Group 2 (Habitats; post-topic)

P1	Animals breathe oxygen		
P2	So do we don't we?		
P3	Creatures breathe oxygen		
.			
.			
P2	Oh there's another one for trees	[Introducing]	
.			
.			
P2	Here's another one for trees, trees what-sha-ma-call-it	[Elaborating]	
P1	Trees what-sha-ma-call-it?	[Contextual]	Making fun
P3	Trees what-sha-ma-call-it trees	[Contextual]	

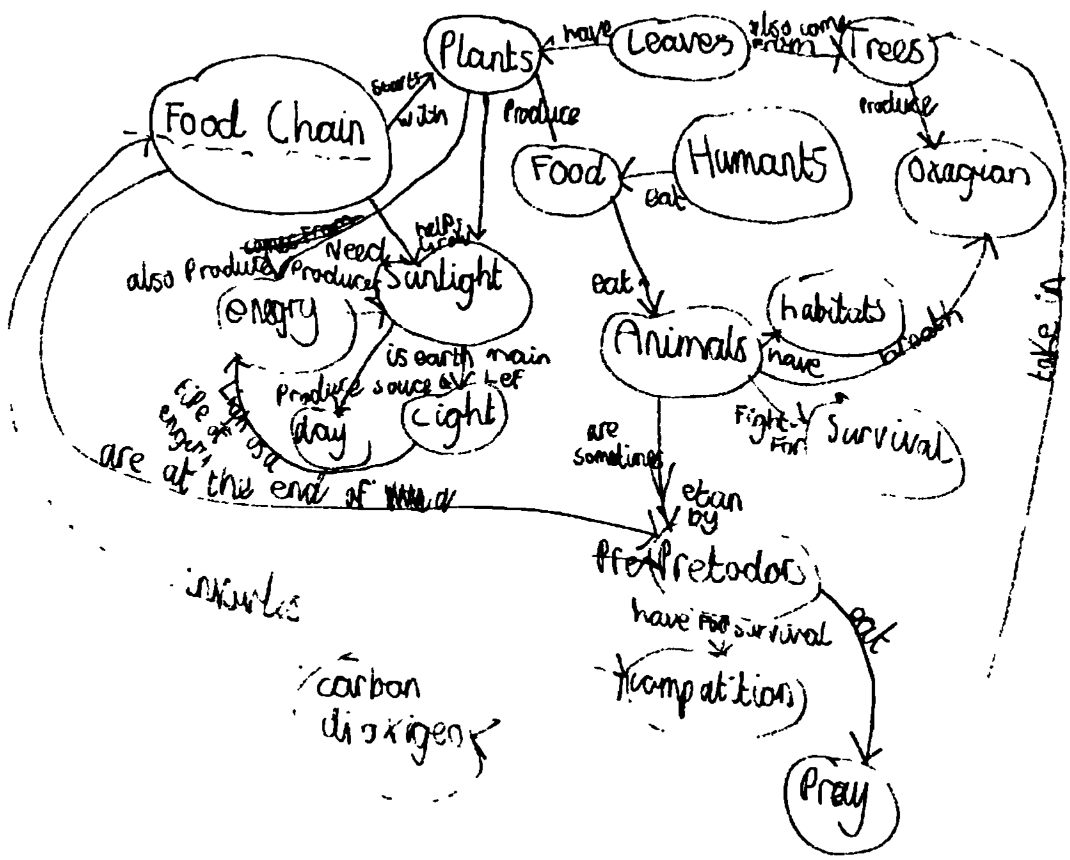
P2	No, trees, um	[Supporting]
P3	Rain	[Elaborating]
P2	What's the word?	[Eliciting elaboration]
P3	Rain	[Supporting]
P2	No listen	[Challenging]
P3	Sunlight?	[Elaborating]
P2	Could you listen please?	[Supporting]
	Trees ... em ... take in carbon dioxide	[Elaborating]

The exchange begins with a reference to the developing concept map: "there's another one for trees". Both the prior discussion and its ongoing documentation in the map conjoin to prompt P2 to make this connection. The "word" that P2 subsequently gropes for in this segment of talk is presumably "photosynthesize". The others in the group attempt without success to put in their own elaborations, but this in itself suggests that they are not mere spectators in the exchange. P2 continues to focus on her original thought, and eventually achieves sufficient elaboration for the point to be made. On the basis of this discussion, the group decide to include "carbon dioxide" and "oxygen" in their concept map.

Analysing this episode suggests the following account. Firstly, a connection occurs to P2, presumably as a result of the preceding discussion. P2 is unable to make the nature of this connection immediately explicit, and so makes an opening bid that indicates where the connection could be made but leaves its content open. In her next move, P2 introduces a place holder ("what-sha-ma-call-it") so as to keep the potential link under consideration. This may simply be to allow more time to recall the necessary wording, or it may be intended to provide a further clue so that others can supply an elaboration. The effect is to keep the idea in play long enough for P2 to come up with what is probably an alternative rendition of the intended relationship. The exchange provides thinking space while the idea, originally vaguely recalled, is clarified. This can be seen as an extension of the function fulfilled by type (i) exchanges, allowing exploration of potential connections without the immediate need to specify the exact relationship involved, but then going on to make that relationship explicit.

The concept map produce by these children (Figure 9.15) shows how they see these constructs as being related to each other, and to the other constructs they have been discussing.

Figure 9.15: Post-topic Concept Map on "Habitats" (Group 2)



These children have clearly achieved several significant conceptual linkages. At least some of these are due to the contribution reproduced above, but all are made accessible to the whole group by virtue of their being transcribed onto the concept map. The rôle of light, as energy originating from the sun is indicated in relation to plants' source of food. Trees are identified as producing oxygen, which is breathed by animals. There are also identifiable gaps which, through being displayed in this way, the teacher could choose to address if appropriate. These include the link between animals and carbon dioxide, the rôle of leaves in absorbing light energy, and the production of oxygen by plants other than trees.

The above example shows that type (ii) exchanges could make a significant contribution to the emerging concept map. But to what extent was that contribution *negotiated*? Examining instances across the data set provided evidence that the initial introduction was intended as provisional in many cases. One sign of this was an overt tentativeness to the introductory move, such as when it was given the form of a question, as in the following example. Here, the children have just linked satellite to moon, and are considering the next connection.

Transcript 18: Group 7 (Earth in space; post-topic)

P1	Planets?	[Introducing]	
P1	Yeah, satellite takes pictures of planets ... let's put satellite to planets	[Elaborating]	
P2	Satellite to planets?	[Challenging]	Sounding incredulous
P1	Yeah, satellite takes pictures, of planets		
P2	Takes pictures ... right	[Contextual]	Rehearsing while writing

P1 introduces a vaguely formed connection, which she then clarifies in the following elaboration, confirming her suggestion that this is a suitable link. P2 seems not to have seen the point at first, but eventually shows his acceptance by saying out loud the words he is writing onto the concept map. In the next example, the type (ii) exchange (which begins with the second of the two introducing moves) forms part of a small cluster of exploratory exchanges as the children search for further links to make, knowing that the session is about to come to an end.

Transcript 19: Group 2 (Habitats; post-topic)

P1	We need another for competition		
P2	Why?		
P3	Cos its only got one		
P1	Yeah		
P1	How about ... prey, prey prey	[Introducing]	
P2	Food	[Introducing]	
P3	No what [?]	[Supporting]	
P2	I know, I've got it ... it's competition for food	[Elaborating]	Sounding excited
P1	Prey	[Supporting]	
P3	Maybe not	[Hedging]	Sounding doubtful

This example shows how the form of the activity could structure the discourse. The children are mindful of the need to make as many appropriate connections as they can, and observe that they have made only one link to competition. They begin to consider possibilities. P2 introduces food, but does not yet seem to have a clear idea of what the link should be. Subsequently, her excitement shows that she has seen how the connection should be made. Although the reception is less enthusiastic, this seems, taking into account subsequent remarks, to be due to the difficulty of drawing the link in an overcrowded map. Eventually, the relationship does get incorporated. The next example shows a similar line of development to that in the previous example.

Transcript 20: Group 14 (Sound and hearing; post-topic)

P1	Um, well, left we got vibrate and guitar string	[Introducing]
P1	Well guitar string and vibrate go together, we know that, cos when you strum a guitar string ... then ... then it, erm ... vibrates	[Elaborating]
P2	Vibrates	[Supporting]

Again, the connection, prompted by the terms that have not been included yet, is introduced in an indeterminate way and elaborated subsequently by the same child.

Being few in number compared with the other types of exchange, there emerged no clear pattern regarding the outcome in the concept map from type (ii) exchanges. Across all the groups, type (ii) exchanges were more likely to result in a scientifically acceptable outcome than type (i) exchanges (Figure 9.9). The implication would seem to be that elaboration over several moves had a positive influence on the final outcome. Post-topic, there was an even greater proportion of scientifically acceptable outcomes than pre-topic. This trend, however, was common to all exchange types, presumably as a result of learning by the children. So although they often functioned in a similar exploratory way to type (i) exchanges, type (ii) exchanges were more often used to develop ideas that were taken forward.

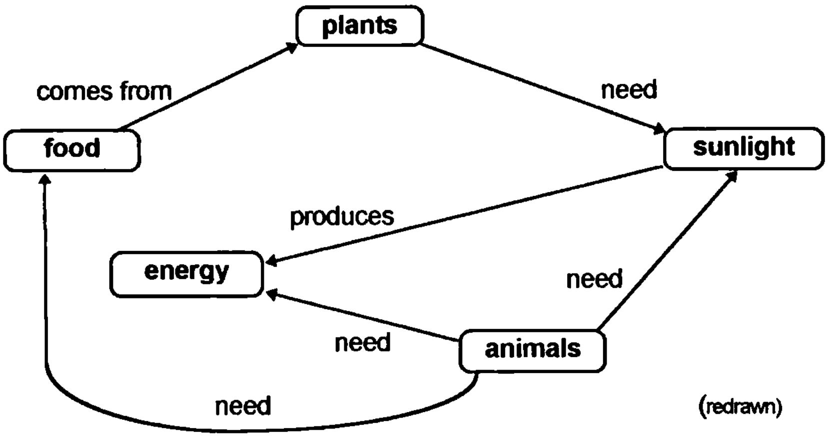
Whilst type (ii) exchanges were primarily one person's contribution to the discussion, this does not imply that other group members were involved at only a minimal level. As with non-elaborated exchanges, the support of the rest of the children determined whether the ideas presented were retained, rejected, or indeed modified. A contribution presented as unproblematic could have its status redefined by somebody asking for further information or for justification. The next example shows a lengthy type (ii) exchange in which a proposition is introduced, made more explicit and justified, all by one child in the group.

Transcript 21: Group 3 (Habitats; post-topic)

P1	Like something needs sunlight to survive	[Introducing]	
P2	Yeah	[Supporting]	
P1	Needs sunlight to survive	[Supporting]	
.			They stray from this exchange for a time
.			
.			
P1	'course there's sunlight, um ... animals need sunlight to survive	[Elaborating]	
P3	Yeah		
P1	Sunlight does need, animals do need sunlight	[Supporting]	
P3	To survive and to live	[Supporting]	
P1	Survive and to live		
.			
.			
P4	Why do animals need it?	[Eliciting elaboration]	
P1	'cos if there wasn't any sun they would just freeze to death, wouldn't they?	[Elaborating]	

In the early stages, P1's repetition of the initial idea serves to keep it under consideration, and he seems to be confirming in his own thinking that it is worth including in the map. Eventually one of the others reveals that she is unsure as to what P1 has in mind regarding this relationship. P1's answer shows that he has some grasp of the importance of sunlight, but he does not go on to identify other ways in which sunlight is vital. However, the concept map that emerges (Figure 9.16) shows that many of the links are in place ready to support this understanding, having been developed at other stages in the discussion.

Figure 9.16: Excerpt from Post-topic Concept Map (Group 3)



This examination of type (ii) exchanges builds on the understandings developed from examination of type (i) exchanges. It shows the utterances of individual children, the distribution of control over the discourse amongst the group members and the transcription of the results of the discussion onto the concept map to be related intimately to one another at specific points in the discussion. Individual children made connections on the basis of what they saw, either already in the map, or amongst the list of terms waiting to be used. Innovations were subject to selection by the others in the group, who approved them, rejected them or asked for them to be improved.

Summary and conclusions

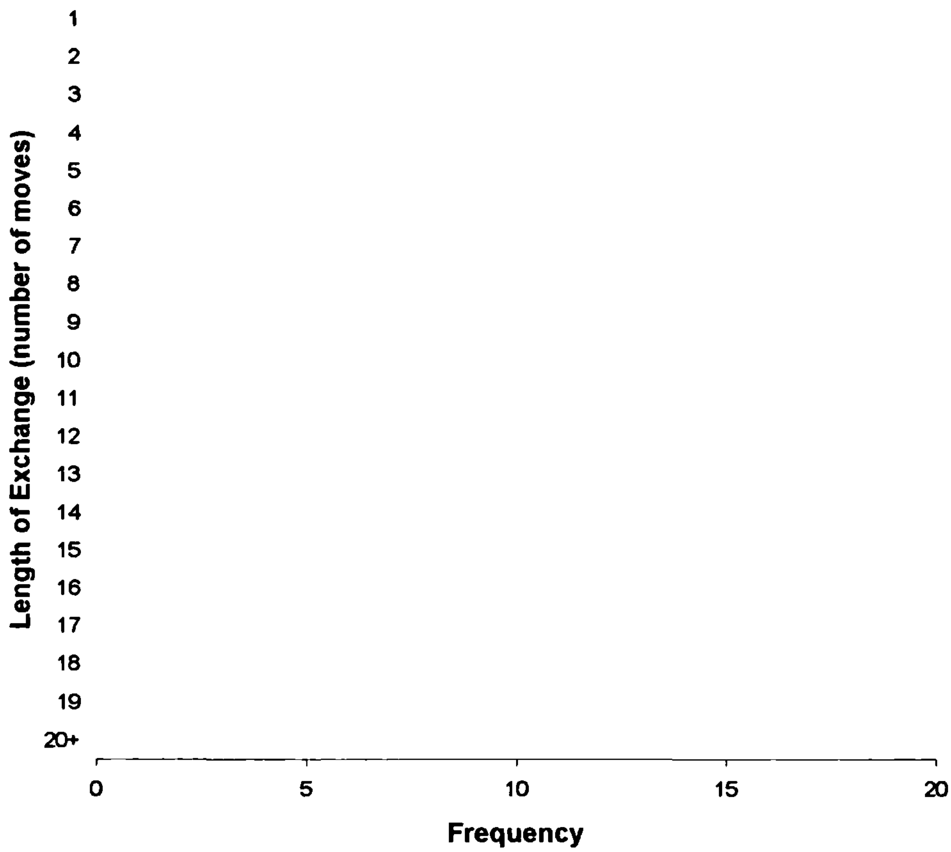
Individually elaborated exchanges were a means of extending the exploratory choices within the discourse. They allowed the children to introduce ideas without waiting to consider fully the nature of the relationships entailed. They gave, in effect, thinking space.

The ideas introduced originated out of the possibilities left open by the ongoing concept mapping activity. Subsequently, they were elaborated upon by their originator, clarifying the nature of the implied relationship. This elaboration took place either voluntarily or it was elicited by the other group members. Frequently, the latter played an active rôle in selecting ideas either for further development or for rejection.

9.5.3 Type (iii) Exchanges

Exchanges of type (iii), (introduction → elaborated → collaboratively) occurred with a similar frequency to type (i) exchanges (Figure 9.4), and therefore are an important feature of the data. They varied widely in length, from two moves up to a maximum of 90. Figure 9.17 shows the distribution.

Figure 9.17: The length of Type (iii) Exchanges



Mean length: 9.36   Median: 8.00   Std Dev: 5.49

Based on data for 152 exchanges

With these exchanges, not only was there the possibility of group approval or disapproval, but also of a direct contribution to the content of the idea under discussion. As with the other two types of exchange, type (iii) exchanges could be introduced with varying degrees of explicitness, and there were indications that they also were being used to introduce or develop ideas in a provisional way. In the first example, the children are just beginning on their map.

Transcript 22: Group 3 (Habitats; pre-topic)

P1	Which one are we going to do first?	[Opening]
P2	I don't know ... em ... animals?	[Introducing]
P3	Yeah	[Supporting]
P?	Hmm	[Supporting]
P3	That is the main one	[Elaborating]
P1	Alright then	[Supporting]
P4	Em ... animals need food	[Elaborating]
P2	Just do animals need food	[Supporting]
P1	For their energy	[Elaborating]
P2	Yeah	[Supporting]
P3	No hold on	[Hedging]
P2	Do animals ... animals need ... animals need food	[Supporting]
P2	Animals ... animals need food	[Supporting]

P2's initial suggestion is put in the form of a question, emphasizing its provisional character. It is approved. P3 then does something that occurs in numerous type (iii) exchanges: she *provides a reason* for supporting this idea, by explicating why "animals" should be put first. The original idea is then extended by P4, and further elaborated in turn by P1. P3's hedge comes to nothing, and the idea is accepted and transcribed.

This example illustrates two important features of type (iii) exchanges. First, there is the joint construction of the proposition written on the map. Then there is the elaborating move by P3 that seems to be saying, in effect "yes, we are on the same wavelength". Intersubjectivity is not only implied tacitly by the support of the various group members, it is actively confirmed through the two forms of elaboration. By extending what they assume to be the originator's intended meaning, the others are able, not only to settle on a link to write in the map, but at the same time to check that these assumptions were correct. From this point of view, P2's initial suggestion can also be seen as a means of checking for intersubjectivity. By putting her suggestion in a vague and tentative way, she can then see whether the reaction confirms her idea or calls it into question. The following two examples illustrate this process at work in different groups.

Transcript 23: Group 9 (Earth in space; post-topic)

P1	Which one haven't we been able to do, what we done ... satellite	[Introducing]	
P2	Satellite	[Supporting]	
P3	Hold on a second	[Contextual]	
P4	Can moon be a satellite?	[Elaborating]	
P3	Moon?	[Challenging]	
P2	M, moon, yeah moon ... moon is a satellite	[Supporting]	Hesitant at first, then confident
P1	Yeah	[Supporting]	
P2	Moon, now is a, innit	[Supporting]	
P4	Yeah	[Supporting]	

P3 seems not to share P4's tentative proposal that the moon might be a satellite. P2's initial hesitance shows that he is thinking through what has been suggested, and then his increased confidence indicates both that he has grasped the point being made and that he agrees with it. P4 eventually gives support as well.

Transcript 24: Group 11 (Sound and hearing; pre-topic)

P1	Ear drum	[Introducing]
P2	Ear drum ... em	[Supporting]
P3	Sound wave ... yeah, it can, it can go like this ... yeah, go on, what was you going to?	[Elaborating]
P2	[?]	[Unclear]
P2	To the sound wave ... because, em right, you can hear sounds, from your ear drum	[Elaborating]
P3	Yeah	[Supporting]
P1	Yeah	[Supporting]

In each of these extracts, there is a sense of convergence on a shared understanding, a negotiation of meaning, that is due to the contributions of more than one participant. This is not convergence from opposing perspectives, and there is no indication that the understandings that emerge are substantively new. There is, nevertheless, a testing out of what is taken (provisionally) to be common understanding. This is distinct from the way the discourse progresses in the other types of exchange. In type (i) and (ii) exchanges, the extent to which members of the group other than the initiator actively check whether understanding is shared is very limited. Essentially, they can agree or disagree, or ask for more information. Through type (iii) exchanges, they can make inferences about what is being meant, about which particular way of using the terms is implied. They can then deduce consequences of that meaning in terms of propositions that would follow from it, and put those propositions to the group for confirmation. As was explained in 8.2.4, meaning depends on inexplicit but shared "constitutive" rules (Searle, 1969), and these rules exist only in their application. By engaging in trial attempts to apply these rules in interaction, the participants are simultaneously negotiating and developing the rules. In some cases, as in the following, a subtle readjustment of the pupils' perspectives can be discerned in the course of this negotiation.

Transcript 25: Group 7 (Earth in space; post-topic)

P1	The moon ... the moon goes round the Earth	[Introducing]
P2	The moon ... orbits the sun	[Elaborating]
P2	Or, orbits the Earth	[Elaborating]

P2 picks up on P1's original suggestion (the moon), bringing to the discussion the more precise term "orbits". But P2 has got the spatial relationship wrong. Hearing from P1 what this should be, P2 then revises the former elaboration, leaving a scientifically appropriate proposition which is then transferred to the map. It is not clear whether P2 actually thought the moon orbited the sun, or whether this was just a slip. But either way, the correct meaning is established.

In other cases, uncertainty leads to a more open request for help in understanding a term or connection. The next example illustrates two further points. Firstly, it shows that not only did the children use trial elaborations to check on shared understanding, they also employed direct questions to elicit the meanings being applied by others. In this case, the question is directed at the teacher.

Transcript 26: Group 11 (Sound and hearing; post-topic)

P1	What's this mean?	[Introducing]	Referring to <cochlea>
P1	Co, co ... co	[Contextual]	Trying to read it
P1	What's this mean?	[Supporting]	
Te	Oh dear, no one's remembered that have they?	[Contextual]	
P2	Oh I know what it is, it's the bone from the ear	[Elaborating]	
P3	What pitch?	[Query]	
P2	Ye, no, cochlea	[Answer]	
P2	The ear drum	[Elaborating]	
P3	Hey, yeah	[Supporting]	
P2	It's an ear drum, ain't it? yeah, right	[Supporting]	

P1 raises for the group the difficulty they are experiencing in attaching a meaning to "cochlea". Eventually, P2 recalls (only partially) some information that is relevant to answering this request. The teacher has moved on by this time. A connection is then written onto the map: "cochlea is part of the ear drum". Later the teacher returns, and sees what they have written.

Te	It's not a part of the ear drum, it's part of your, ear	[Elaborating]	

To illustrate this point, the teacher directs them to a model ear, which is on display in the classroom. However, they still regard the cochlea as being a bone (the model could hardly refute this), and describe it as such in their map.

This example underlines the developing theme. Within the groups, there was a striving to make sense of the construct terms they had been given, and how these are related. However, they did not always have the means within the group to sort out their difficulties. This is an important point, and the subject of further discussion in 9.8. In this present context, we note that the concept map served to document their decision. This document was then available to the teacher, and helped to highlight gaps in the children's understanding. This in turn enabled the teacher to take action to bring about improved understanding, albeit that in this example the action was not completely effective.

As with the other types of exchange, there were examples of how the physical presence of the construct terms, together with continuing dialogue concerning the relationship between those terms, was inspiring the children to consider potential new links. In the following example, this process is exemplified.

Transcript 27: Group 1 (Habitats; post-topic)

P1	En, ener ... food comes from	[Introducing]
P2	What do animals [?]	[Eliciting elaboration]
P1	In a habitat ... forests ... woods	[Elaborating]
P1	Food chain	[Introducing]
P3	Food chain and habitat, cos there might be a food chain in a certain habitat	[Elaborating]
P2	Mmm ... could be ... yes because of where they live, that could be a ... that could be food whether animal or bush	[Elaborating]

P1's first introducing move here is tentative and exploratory. It seems to prompt a question from P2, although what is being asked is not entirely clear from the recording. P1's next move must be taken as a reply to P2, as it is incomplete without seeing it as a response. The term "habitat" is exemplified as "forests ... woods". Next P1 utters the single term "food chain", and P3 comes into the discussion with an elaboration of the connection between this and what they have just been discussing. P2 confirms that the connection makes sense.

In the following example, the children are introduced by the teacher to the term "decibels", which they have not yet met. Following an explanation by the teacher, one of the children makes a connection to one of the terms already located in the concept map.

Transcript 28: Group 10 (Sound and hearing; pre-topic)

Te	This is, this is decibels ... and decibels, are the measurement of loudness	[Introducing]	
P1	What, like volume?	[Elaborating]	
P2	Loudness?	[Query]	
Te	Volume	[Supporting]	
P3	Sound waves, sound waves	[Integrating]	Very excited
P2	Sound waves	[Supporting]	
Te	Well done	[Supporting]	

P1 is the first to attempt an elaboration building on what the teacher has said, and this is put in the form of a question inviting feedback on whether shared understanding has been achieved. This is duly given, but in the mean time, P3 has noticed another connection. His excited tone indicates that this is, for him, a significant insight: decibels are connected to sound waves, because "loudness" is a property of sound waves. The teacher confirms that there is a valid connection ("Well done"), although this is not elaborated further and the other children continue to focus on the definition of decibels.

This last example seems to be a case in which a genuine insight occurred, a new connection that the children were not aware of before the session. Although other examples occurred, it was rarely the case that entirely new understanding was achieved. Far more common was the process of negotiated adjustment described earlier. The negotiation and development of scientific meanings is explored further in 9.7.

Summary and conclusions

Collaboratively elaborated exchanges were the main vehicle by which scientifically acceptable relationships were negotiated for incorporation into the concept map. In these exchanges, children asked for, and gave, explanations of the connections and terms being discussed. They also had the opportunity to probe in more subtle ways their assumptions of shared understanding of the idea being discussed. By adding to ideas introduced by others, the children could test out whether they were applying the same set of meanings. Through this process of hypothesis testing, the children were able to confirm those assumptions or to adjust their perspective according to the outcome. It was rare for this to result in a major change to shared understanding. However, small readjustments of meaning were in evidence. The terms being mapped affected directly the extent to which scientific meanings were negotiated in producing these concept maps.

9.6 A Distinctive Form of Discourse

It is now apposite to weave together some of the threads of evidence that may be drawn out of the preceding analyses.

The first point worthy of note is that the form of dialogue is quite different from that found in the typical "triadic" teaching exchange of question/answer/evaluation (Lemke, 1990). In triadic exchanges, the evaluation move is used to compare the answer given to the one expected. In the discussions analysed here, there do not seem to be any expected answers, and there is no direct equivalent of the teacher's evaluation. Instead, there can be a range of different responses to a suggestion, varying from a challenge, through non-committal hedging, to support, and then, most significantly, to moves that elaborate on the initial response.

These moves are evaluative in some senses. Firstly, they indicate acceptance, uncertainty or rejection. But this is evaluation against the developing set of meanings that forms part of the task rather than against any predetermined result. Also, these moves can serve to evaluate the extent of common understanding and intersubjectivity within the group, through testing provisional meanings. Sequences of talk are patterned by the nature of the power relationship between the interlocutors, as well as by the subject matter of the discourse (see, for example, Willes, 1983). The process of trying out interpretations and adjusting them is encouraged in a context in which tentative moves are acceptable, and therefore in which the knowledge of one person is not valued far in excess of the others. Teaching exchanges characterized by the familiar opening, answering, follow-up of teacher-class interaction simultaneously embody and maintain the teacher-class power relationship. In such cases, it is the teacher who decides what is relevant, and what it is legitimate to say. This point is well-recognized by Edwards & Mercer (1987). In contrast, within these pupil groups, the authority to define what is relevant was essentially delegated, and distributed amongst the group members. Hence the discourse took on a different and more exploratory character.

One of the most significant features of this distinctive form of discourse is the phenomenon of children's finishing off each other's contributions. This often occurred over successive moves in the course of type (iii) exchanges, as was discussed above. It also happened that children completed individual utterances begun by others, so producing jointly constructed discourse move. Roth & Roychoudhury (1993) take this to be a sign of intersubjectivity in the groups they studied, since to take up and complete the thoughts of another must involve assumptions of shared understanding. The following are examples of jointly constructed moves from the present data. In the first, P1 begins the move, and P2 completes it almost simultaneously with P1.

P1	So do satellite, takes pictures of, and then ... um, planets
P2	Planets

In the second example, from a pre-topic session on sounds, P1 starts off a proposition, only to have it completed by more than one of the others in chorus.

P1	And gu, guitars string
Chor.	Vibrates

The third example shows the development of a proposition over three turns:

P1	We could put habitat here ... and then we could put animals
P2	Have different
P1	Habitat, yeah

Examples of this phenomenon, though infrequent, are scattered throughout the data. There is at least one instance in all except one of the transcripts. It would be foolish to suggest that they form any sort of metric by which to measure intersubjectivity; they are too dependent on chance factors for that (see also Barnes & Todd, 1977, who ultimately rejected the possibility of such metrics). But they do serve as further evidence that the groups were working collaboratively towards shared meaning, and indeed the single negative case was for group 5 (post-topic discussion), a pair of children who did not engage in very much discussion of meaning at all.

Another significant phenomenon is illustrated by the example given above in Transcript 28: it is one of the pupils asking the question and not the teacher. This can also be seen in other transcripts. In Transcript 26, there is another example of a question to the teacher about the meaning of a term. In Transcript 23, there is a question to the group. These are not pseudo-questions or assessment questions, as used particularly by teachers to elicit information from pupils that they already possess themselves. They are genuine enquiries to obtain new information, and are another of the means used to move towards shared understanding. And it is shared understanding that gives content to the inscription in the concept map. The preceding analyses have shown that the various processes of joint construction and negotiation of meaning are operative in keeping that content focused on scientific meanings.

Concept mapping has been presented in this thesis as a means of developing children's understanding of the relationships between constructs in a scientific domain. The findings reported above would support this claim. Relationships were certainly being considered, discussed, clarified, built up and evaluated in the course of the discussion, and drawing on contributions from across the group. This in itself suggests that the activity is of value, a point that will be developed in the final chapter. But in evaluating the full potential of the discussion to contribute to children's understanding of science, some further questions remain to be investigated.

- *Did the discussion provide opportunities for shifts in understanding, from misconceptions and everyday meanings towards scientific meanings?*
- *Did the discussion involve children in addressing epistemic questions concerning the evidence for the relationships they were discussing?*
- *What links were there between the concept mapping activity and the wider context of scientific information and activity available to the children?*

9.7 Developing Scientific Meanings

In this part of the chapter, examples will be presented that show shifts in the ideas under discussion from everyday usage and misconceptions towards more scientifically acceptable meanings. This is best done in relation to the three topic areas, so that instances of specific relevant everyday meanings and misconceptions can be considered.

9.7.1 Habitats

There were few misconceptions arising in this topic area (see Figure 9.6), and the main area of growth was therefore in the movement from everyday to scientific meanings. Terms that were sometimes being used in their everyday sense included "energy", "competition" and "animal". Energy was seen as closely connected with activity, and competition as primarily a conflict between individual animals, a more concrete idea than the scientific usage. These views were reflected in the following quotations.

> Sunlight gives energy, to animals and plants, and humans

Humans are seen here as distinct from animals.

> Food makes energy ... and animals need energy

Energy is seen as a product that is necessary (for activity) on the part of animals.

> You need energy for competition

Although not incorrect, this last example inverts the scientific emphasis that organisms compete for resources, including energy supplies.

Another area in which the children's grasp of scientific meanings was limited was in the nature of the hierarchy of ideas. It was common for the children to place animals at the centre of the network of relationships (see Transcript 22). This gave animals a higher status than plants, which were seen primarily as a food supply rather than as organisms in their own right. These various conceptualizations would presumably have their origin in an essentially egocentric position, in which biological knowledge develops out of a "naïve psychology" framed initially in terms of human behaviour (Carey, 1985). However, there were occasions in which these everyday meanings were transcended, and elements of a more scientific perspective introduced. Some examples have already been presented and commented upon in the previous part of the chapter. Others will now be discussed.

One of these occurred with group 2 in the post-topic session. An initial move to put "animals" as the major idea was abandoned in favour of "food chain", although there was no explicit reasoning in support of this. Following this, animals and plants were treated more even-handedly. The culmination was the consideration, as in Transcript 17 and Figure 9.15 above, of the interrelatedness of plant and animal respiration.

The next example involves the meaning of "competition".

Transcript 29: Group 1 (Habitats; post-topic)

P1	Predators ... attack animals ... predators attack animals ... food chain
P2	Hmm, I'm not sure about predators
P3	Predators is food chain cos they it eats the an, animals
P1	Yeah cos it's like they're all linked together
P3	Mmm
.	
.	
P1	Animals, no predators attack other animals ... so predators attack other predators
P1	Competition ... competition between the predators
P1	Will competition be ... will the competition be, between a predator ... or a food chain?
P2	Both
P3	Both

P1 starts off from the idea that predators attack animals, and seems to be using the term "animals" to stand for "prey" in this instance. In fact, there is a general failure to distinguish animal, prey and predator, at various steps in this segment of talk. P2 has not grasped why predators link with food chain, and, although it is not clear from what is said, this may be due to P2's seeing that "animals" includes predators. P3 and P1 press on with the link, and P2 tacitly allows this. P1 then observes that if predators attack animals then they are attacking other predators, apparently conflating animals and predators together. However, this leads P1 to make the connection (in an excited tone) that this implies a sort of competition. This seems to be seen at this stage as a competition between individuals. However, P1 goes on to question whether there might be competition between entire food chains. Later, they consider the link between a food chain and a particular habitat (see Transcript 27). Through these moves, the group edges slightly further on in their grasp of the interrelationships between different animals in a habitat, which in practice are more complex than a simple food chain. This discussion has certainly not sorted these ideas out fully. But it has just as certainly progressed in terms of the relationships under consideration, due to P1's inspired introduction of "competition".

The above example points up a general area of ambiguity in these concept mapping sessions: shifting reference. An example will establish the point. The twin propositions "animals eat plants" and "animals eat other animals" may each be either true or false, depending on what precisely is being referred to. Using "animals" as a general term, the relationships are invalid: not all animals eat plants, and not all animals eat other animals. On the other hand, the statements may be taken as elliptical for "certain individual animals eat plants" and "certain individual animals eat other animals", as, for example, in response to the question "What is it that eats plants?" When the starting point is a word on a slip of paper, then this can be, and apparently was, taken in different ways in the course of the discussion (and possibly, by different people, as appears to happen in the preceding example). This is a sign of the activity's becoming decontextualized, as a task in its own right, and losing its links to the wider context of scientific activity and knowledge. It also occurred with respect to "humans". Humans as a species are predators. But not all humans are carnivorous. Consequently, in some of the groups, a disagreement arose as to whether humans should be classed as predators, because the children could think of exceptions. This element of ambiguity was double-edged. In some cases it was skipped over, and the opportunity to clarify meaning was missed. In other instances, the ambiguity allowed new possibilities to open up. In the preceding example, both are illustrated: the unresolved conflation of animal and predator and the serendipitous link to competition, which is elaborated upon subsequently. The next case shows how ambiguity between the everyday notion of animal and the scientific construct led to greater clarity of meaning.

Transcript 30: Group 2 (Habitats; post-topic)

P1	What we can do, is do predator here, ... then we can do animals eaten by predators
P1	Predators can be animals ... predators are animals, aren't they?
P2	But not always
P4	No, predators are animals
P2	Not all of the time
P3	No because
P2	Like humans
P4	We're animals
P3	We're predators ... we're predators
P1	Yeah, yeah, but we're still animals though ... and we're herbivores
P3	Do it omnivore
P4	What's that mean?
P3	Mixture

P2 responds to P1's claim that predators are animals by claiming this is not always so. He cites humans as an example (of a predator that is not an animal). P3 sides with P2, making explicit what P2 has left implicit: that humans are predators. P4, in the other hand, expresses the scientific view that humans are animals. P1 is quick to reinforce this, adding on the way that humans are not exclusively carnivorous. P3 falls in line, supplying and briefly explaining a term (omnivore) to denote this.

These extracts show that progress towards scientific meanings did take place. But the steps made were small. There were also clearly times when some negotiation of meaning would have been beneficial, but was not attempted, as in Transcript 15. The topic covered by these concept mapping sessions was complex, and it is unreasonable to expect these discussions to bring about large advances in understanding in isolation. Yet there does seem to have been a high degree of isolation associated with these tasks. At no time did the children refer explicitly to other learning experiences in order to fix the reference for one of the construct terms (although it might nevertheless have been the case that some of what was said was influenced by those experiences). The nature of the relationship between concept maps and the wider realm of experience is an important issue, and is taken up in 9.8. Meanwhile, consideration of the progress made in the different topic areas continues in the next section.

9.7.2 Earth in Space

In contrast to the previous one, this topic was resplendent with misconceptions as well as everyday meanings that overlapped with scientific constructs. In terms of everyday talk, both "moon" and "satellite" have a more restricted meaning than in scientific discourse. Moon is taken to apply to the one that can be seen at night from the Earth, and the moons of other planets are understood by projecting from this sense. Satellite denotes a manufactured article, a "sputnik", and does not serve as a general term for an orbiting object. Several misconceptions were encountered in addition to these everyday meanings:

- the sun disappears behind a cloud at night;
- planets (or Earth) are the same as stars;
- the universe is an object;

- the moon is a planet;
- the sun orbits the Earth;
- the stars orbit planets (or the Earth).

These all arose during the course of discussion, but it should not necessarily be assumed from this that what arose had the status of established ideas; some probably appeared by way of speculation. Nevertheless, that they appeared at all is significant, and how they were dealt with is consequently of interest.

The first extract shows how one group of children moved towards a theory in which moons are common to planets other than Earth.

Transcript 31: Group 7 (Earth in Space; Pre-topic)

P1	Oi, connect it up with moon and planets and all ... 'cos all planets have moons
P2	[?] what is it?
P3	Say that again?
P2	The sun is the
P3	Say that again about the moon and the planets?
P1	'Cos all, most of the planets have a moon, so if we connect
P4	The moon to, to planets?
P1	Yeah

It is clear from the tone of the children's reactions that this is a new idea to the other members of the group. They do not challenge it, though, and it is incorporated in the group's concept map. What is even more significant is that the same link emerges in the group's post-topic discussion, from which P1 here was absent. In the latter session, the idea is introduced and readily agreed to, and the link is drawn and labelled on the map. The teacher's planning notes included a model concept map of the topic; this did not feature any direct link between "moon" and "planets". Also, none of the other groups considered the construct "moon" in this way. So this may well be evidence of learning by the children in group 7 that is attributable directly to the pre-topic discussion, rather than to other learning experiences provided by the teacher.

It is the "sputnik" meaning of "satellite" that was adopted by most of the children during the course of this topic. However, in one group one of the children attempted to exchange this meaning for a more scientific one.

Transcript 32: Group 9 (Earth in Space; pre-topic)

P1	Then you can do, the moon is a satellite
P2	Satellite?
	[laughter]
P1	It is a satellite
	[laughter]
CA[22]	The moon is a satellite, ***
P3	I know, I never said it wasn't, it isn't quite how you meant
P1	I don't mean the metal sort of satellite, I mean

Eventually, this is accepted. As with the previous example, the same idea emerges again in the post-topic concept mapping session. And similarly, the pupil who originally introduced the scientific meaning was absent from the group on this second occasion. Once more, this is evidence that children learned from the sharing of ideas.

In group 7's pre-topic session, one of the group introduced the scientific meaning for "satellite":

> Yeah but a satellite could be a big piece of rock, that can be a satellite, as well as a little bit of tin that gives out signals

However, in this case the idea did not find it's way into the concept map, and neither did it re-emerge in the post-topic discussion. Its significance does not seem to have been realised.

In only one other instance was there an indication of an everyday meaning's being introduced and ultimately displaced by more scientific language. This occurred in group 9's pre-topic session when the classroom assistant introduced a geocentric viewpoint:

> The Earth has a sun and a moon

This was not challenged directly, though the discussion led on to a more scientific view, in which the Earth is seen as rotating about the sun, rather than as "having" a sun. In this case, then, the vague idea initially introduced acted as a stepping-off point for a more specific formulation.

It was clear that in the early, pre-topic, session, many of the children were not sure of the nature of stars, or of how they differed from planets. Consequently, when children suggested that the sun was a planet, this

---

[22] CA = classroom assistant (see Appendix E for a full explanation of the symbols)

typically went unchallenged, even when the children knew (at least by rote, if not with any depth of understanding) that the sun was a star. This extended in some cases to the post-topic session as well. One of the transcripts for the post-topic session featured a long exchange, over the course of which the children debated whether the sun was a planet. The question had arisen, been set aside, and reappeared twice before one of the group attempted to put together an argument based on the idea that it is not possible to visit a star, while it is possible to visit a planet. However, the pupil proposing this argument did not appear to know why it is not possible to visit a star. One of the other children soon spotted the weakness in the argument. In the following extract, she takes up the challenge.

Transcript 33: Group 6 (Earth in Space; post-topic)

P1	Say we go to Jupiter, right? ... and that's a star before we go there, and we land, well, and we, um, somebody steps on it, so that's not going to be a star any more, it's going to be a planet?
P2	Yeah
P2	Yeah
Te	What's Jupiter?
P2	It's a planet
P3	Is Jupiter a planet?
P1	Yeah, but how, how's it a planet if you don't, if you, if you can't go, go, go to Jupiter?
P3	I'm not talking about you can't go to it
P2	Yeah but Jupiter you can see it
P1	Yeah but you said you can, um
P?	Yeah, but on a satellite
P1	You can't, you can't go to the sun, and you can't go to the stars
P?	No, stars
P1	So they're stars ... if you can't go to Jupiter, it's a star then, isn't it? Cos you can't go to the sun

Eventually, the argument reaches an impasse, as the group lack the understanding to resolve the issue. They decide not to make the link on their map, but this seems to be adopted as the easy solution to the disagreement, rather than a definitive decision. It is possible that such an incident would serve to highlight for the children their lack of understanding, and so create an awareness of the need for further learning. However, there is no way of knowing whether this happened in the case in question. In the next extract, on the other hand, the participants acknowledge openly that they do not have the understanding to answer a challenge that is thrown up, and consciously set it aside for further consideration.

Transcript 34: Group 9 (Earth in space; pre topic)

P1	Yeah, but the moon might be, the moon, might be a planet
P2	Well it might but, we can talk about that after

Not all the instances of misconceptions or conflicting conceptions could be viewed positively. In the following segment of talk, the children seem to be working with a notion of universe that is far removed from the scientific construct so far removed, in fact, that it is difficult to understand just what sort of notion it is that is being entertained.

Transcript 35: Group 8 (Earth in space; pre-topic)

P1	OK now, the universe is the big one, so the universe is the
P2	The universe is a big planet
P1	N no ... it's the whole thing
P2	Yes it, yeah, the universe is the whole world
P1	Whole thing and whole world
P2	Yeah, yeah
P1	Now we've got to sort it out, does Earth go to universe ... universe ... is Earth kind of a universe?
P1	Earth is kind of a universe, is it?
P2	Don't know
P1	Planets are universes
P2	Yeah, 'course
P1	Stars are
P2	Stars ain't a universe you idiot
P1	All right then, em, planets, planets, planets is a universe, Earth is a kind of universe

P1, lacking secure knowledge of what ontological category universe falls in, is lead away from the astronomical sense by P2. This seems to be a good example of "the blind leading the blind". The same confusion did not arise in the post-topic session, so there must have been some progress in understanding over the course of the topic. However, this did not seem to be due in any way to the discussion reproduced here.

Only rarely, during these sessions, did the children make explicit reference to experiences outside the activity itself in order to clarify areas of obscure meaning. These sessions too seemed to exist somewhat isolated from other learning experience.

9.7.3 Sound and Hearing

In the third topic area, sound and hearing, a slightly different pattern of conceptions occurred. As with the first two topics, there were everyday meanings for some of the terms, usually taking the form of a vaguer notion than would be desired from a scientific viewpoint. The children tended to think of "echo" as a general term for a certain quality or tone of sound. "Loudness" was seen as a description applying only to loud sounds. Soft sounds were not considered to have loudness. This suggests a general point, that attempts to use everyday terms to stand for specific scientific constructs (volume, in this case) face the difficulty that the everyday term may have a usage that differs in significant respects from the technical term. To balance this, the opposite situation also occurred, in which technical terms were given to the children (pitch, and decibels pre-topic) to which the children could attach no meaning. The term "vibrate" was typically seen as a correlate of a sound, but not causal.

The only misconception to appear was when one child seemed to think that "pitch" referred to a hollow sound quality. This was more likely a lapse of memory than an established alternative meaning.

Some of the groups spent time considering which term was most important in the map, as in the first extract.

Transcript 36: Group 10 (Sound & hearing; post-topic)

P1 Right, who, shall we have a vote and see who wants vibrate at the top?
P2 Sound wave might be important
P3 <u>Yeah, it might be</u>
P1 <u>No ...</u> no but ... no but, yeah but, how, how do you make sounds, it's with vibration, isn't it?
P2 <u>Not all the time</u>
P1 <u>Vibrates and makes sound</u>
P4 It vibrates the little ... what is it
P1 Yeah, plus
P4 Molecules in the air, it vibrates the molecules in the air, which makes the sound, waves travel to the ear ... drum

P5 You know what we saw about the sugar and the drum ... the vibration
P? Yeah
P2 The air?
P4 Yeah so the air molecules that makes the sound waves travel
P1 And, and, you couldn't hear with your ear drum, if you never had a vibration
P2 All right ... put vibration

In contrast to the common view that vibration is a correlate of sound, in this segment of talk the children adopt a position in which vibration has wider explanatory power. P1 indicates that vibrations lie behind sound waves, and P4 elaborates to explain in more detail how vibration propagates through the air. Meanwhile, P5 introduces evidence from a shared experience, that of watching particles moving on a beaten drum skin, which backs up the claims about the connection between vibration and sound. P2 accepts the arguments made, and agrees to the importance of vibration. In this discussion, then, several threads are joined to create a more unified and scientific view of sound, threads that include both theoretical entities (molecules) and practical experience (the drum).

Group 13's pre-topic session included a discussion that focused on the nature of echoes, another construct that tended to be employed in a vague way by many of the children.

Transcript 37: Group 13 (Sound & hearing; pre-topic)

P1 Sound waves ... sound waves bounce off walls creating echoes, OK, so they, they definitely go together
P2 In big caves, in caves, in caves
P1 No
P3 They bounce off anywhere
P2 In caves
P1 Sound waves bounce off walls, creating echoes
P4 They don't bounce off anywhere, I'm not with you [?]
P1 They don't bounce off the air, 'cos like, if your in the middle of the sea, and you shout, help, then you're not going to hear your echo, are they, cos there's not a wall two feet next to you
:
P2 Sound wave is created
P1 When an echo
P1 No
P2 No
P1 The sound wave's bouncing off walls, make echoes

Here the group establish that an obstacle is needed, from which sound waves are deflected, for there to be an echo. The meaning for echo that emerges therefore approaches the scientific view. This view also emerged in the pre-topic discussion for group 14. The idea had been expressed that an echo was a sound wave (because you hear sound waves), which was elaborated upon by one child who said:

> Yeah, it's a sort of double wave

by which he seemed to mean a reflected sound wave.

The recordings made during the third topic featured more, and longer, discussion between the teacher or researcher and the pupils than did either of the other two topics. Some of these occasions were marked by progress towards scientific meaning, as was the case here. In this next extract, the children pursue with the researcher the idea that there might be a connection between "pitch" and "loudness".

Transcript 38: Group 13 (Sound & hearing; pre-topic)

P1	Can pitch be involved with loudness? Or can there be a quiet pitch?
Re	It might be involved with loudness, but, do you know how? ... What do you mean by pitch?
P2	Like, er, loud 'n pitch ... like you might play a quiet note, it could be a quiet pitch
Re	Is there a difference between loudness and pitch?
P3	Yes
P4	Pitch is like, how high it is, or how low it is
Re	Good, right
P1	Yeah loud can be high
Re	So ... that's not the same thing as how loud it is then, is it?
Re	Is he right then? ... Pitch is, how high or low, a sound is
P3	Suppose so
Re	So what's the difference between that and loudness?
P3	Well, loudness is different volume ... and that's loudness and that's, the tune of it
P2	That's what the note is
P1	That's what it sounds like, and that's how loud it sounds

From the initially vague view of the relationship between pitch and loudness, the children are encouraged to reflect on what they know about these two constructs in order to clarify the distinction between them.

For this topic as well, there were instances of constructive discourse and of movement towards scientific meanings. Unlike the other topics, there were times when the children made reference either to specific experiences, or to material present in the classroom, in order to guide this process of construction. One reason for this was undoubtedly the nature of the terms given to the children to use in their maps. The term "guitar string" denoted an object to be found in the room, and the terms "ear drum" and "cochlea", as parts of the ear, could be related to a model human ear which was in the classroom for the post-topic session. This reference to specific objects in the map and the presence of examples seemed to make it possible for the children to anchor their discussion in ways that were more difficult for concept maps about habitats, plants and animals as generic terms, or about planets, stars and the solar system, which it was not possible for the children to manipulate. The question of resources to support discussion is examined in more detail in the next part of the chapter.

Summary and conclusions

There were occasions when the discussion led the children to clarify or modify their views to conform more to accepted scientific theory. Such occasions were rare, and most of the examples are presented above. Progress in these cases was generally slight. However, these instances would indicate that the discourse was not merely reconstructive. The potential was there to explore new meanings.

Some difficulties were identified in moving towards scientific meanings, that could have implications for the way concept mapping is incorporated into science teaching. Prominent amongst these is the question of links to learning experiences and sources of evidence beyond the concept mapping activity itself. It is to these matters that the discussion now turns.

9.8 Resources for Discussion

In the above examples, progress made towards scientifically more acceptable meanings was incremental. There was nothing to suggest that monumental shifts in understanding occurred. To expect such would be unreasonable. But if shifts in meaning are modest, then it is important that they should impact beyond the limits of the concept mapping activity. The establishment of relational links between the given terms should not be a decontextualized exercise, but ought to engage with further knowledge and action of a scientific nature, which is to say with sources of evidence. This brings us to focus on the wider context in which the discussion took place. Context includes the resources drawn upon to make sense of the terms given to the children.

In addition to the semiotic resources of the language community, three main kinds of resource were identified as feeding into the discussions. These were:

- the explanations or reasoning used by the children based on prior knowledge, and including, for example, using analogy, identifying anomalies and giving examples;
- reference to the material world as personally experienced, and including observing objects and conducting experimental manipulations (either those they have already experienced in the past, or those they conduct in response to questions raised by the discussion);
- the influence of different "voices" from outside the immediate group, including the teacher and reference material such as books.

The uses of these different resources are intertwined. In practice, the teacher or the pupils could, for example make use of objects whilst making a point. Or, in the midst of a discussion in which the children refer to objects or give explanations, the teacher could appropriate the conversation to draw attention to something the children might not otherwise have considered.

To an extent, the availability of these resources is related to the nature of the domain being mapped and to the terms chosen. Reasoning and explanation are dependent on relevant prior knowledge, on the part of the children themselves, to feed into the discussion. But they are at least approaches that can be applied across domains of knowledge. Only in certain aspects of a domain is direct access possible to objects that can be manipulated, and so both further experimentation and examination of materials are restricted to concept maps focusing on these more accessible realms. Other voices, represented by reference material or the teacher, are a type of resource that may vary in accessibility, due perhaps more to the teacher's emphases and provision than to the availability of the resource *per se*. In the classes concerned, for example, use of information books was proscribed during the concept mapping sessions, although the children had the option of recalling materials previously encountered.

There were varied instances in the data of children providing their own explanations or arguments to resolve questions raised in discussion. In one sense, most of the discussion was of this sort. What are focused on in the following examples are cases where constructs other than those included in the concept map were deployed explicitly, rather than as subsidiaries, in order to make a point. Depending on the quality of the prior knowledge represented by these constructs, the results could be more or less successful.

One simple example of this resource in action was exemplification by one of the group members of a term under discussion so as to make its meaning better understood. The first extract shows this happening. It begins with P1 expressing ignorance of the meaning of "food chain".

Transcript 39: Group 3 (Habitats; pre-topic)

P1	Food chain ... I don't know
P2	Well it's sort of like food chain's sort of like when a snail eats a fly or whatever snails eat ... no snail eats a leaf and then a bird eats the snail and goes on like that
P1	Oh, right so like first you start off with a dog ... dog eats cat, cat eats a ... a bird, the bird eats a ... snail, snail eats a leaf

The explanation seems to be completely successful, as evidenced by P2's extension of the example by way of feedback to P1. This is a case of filling out the meaning of the terms provided in the map to obtain clarity, and following that, consensus and progress. In the next extract, a similar process helps the children understand better the nature of sound waves, a construct that seems to have been unfamiliar to some of this group. The teacher is listening in on the discussion.

Transcript 40: Group 10 (Sound & hearing; pre-topic)

P1	Sound waves ... it's like when a, you know dolphins, they have sonar, don't they?
Te	Say it again
P1	Dolph, sound waves, well that's like dolphins, you know when they have, um, sonars, well that's kind of like a sound wave
P2	Yeah, and it goes through
Te	Ex, explain how that, explain how you that, what happens ... when you say

dolphins
P1 When they speak to each other like, it, and, they, can sense, like animals, and things like that
P2 It goes through the water
Te What does?
P2 The noise, it travels
P1 The noise
P1 The sound waves

This exemplification seems to go some way towards clarifying the idea of sound waves, and particularly of sound travelling. Following on from this, the teacher raises a further question for them to consider.

Te Do you think this [vibration] could fit into that idea you're talking about?
P2 You have a vibration ... the noise vibrates through the water, so
P? The what?
P2 The noise vibrates through the water
P3 In sound waves
Te In sound waves?
P3 Mmm
Te OK, so you've got
P3 And through the air
Te And through the air?
P3 Mmm
Te Yeah

In this instance, the example provided by one of the group leads to a definite step forward in negotiating a scientific meaning for the term "sound wave", especially in relation to the rôle of vibration in propagating sound through materials, including through the air. This is a step that owes a great deal to the teacher's well-placed prompting of the children's prior knowledge. The next example shows how an apparently anomalous case can suggest a rethink of the assumptions made up to that point

Transcript 41: Group 4 (Habitats; pre-topic)

P1 Plants are predators
P2 They're not
P3 They're not predators
P4 Venus Flytraps are
P1 Yeah
P3 Yeah but that's not that's not that's a fly plant ... we're not saying fly plant, we're saying plant

This anomaly seems clearly to be one to which the children do not have an immediate answer. Unfortunately, the rethink is not forthcoming. P3's very weak response does not address the issue raised at all, but the group do not pursue the problem any further. The children do not seem to have the resources themselves to resolve the discrepancy, and make no attempt to draw on wider resources. This is a missed opportunity to enrich their understanding of predation.

The following example shows again how the resources of the children's own understanding can be exhausted by conflict, with the result that progress in their understanding does not seem to be made. Part of this exchange was presented above in Transcript 5. In this next extract, the teacher tries to move the children on in their thinking to establish more relevant relationships.

Transcript 42: Group 6 (Earth in space; post-topic)

P1	Yeah but you said you can, um, you can't, you can't go to the sun, and you can't go to the stars, so they're stars ... if you can't go to Jupiter, it's a star then isn't it? Cos you can't go to the sun
P3	It's not impossible to go to Jupiter
Te	Well you're actually having difficulties with the, um, definition ... of what is ... what is it, the definition, of um, a star and a planet?
P2	Well we know that a star's only white, and Jupiter, when we saw it on the satellite pic, when, right, um people, like they look at that, they stay and take pictures, they can see that it's all different colours, so it's bound to be a planet, it can't be a star
P1	Can you see the sun, what colour's the sun?
P?	Yellow
P2	Planets are stars ... stars are planets, ain't they?
Te	If you look at the behaviour, if you look at
P?	Just do it how we did it, yeah?
Te	Go on ... OK, draw the lines in on that then, and see if you're happy with them

One of the children reminds the group of some evidence relevant to the argument: a satellite photograph they have seen. However, it still seems that they lack the underlying understanding of the nature of stars and planets from which to interpret the evidence supplied by the picture. The picture cannot provide them with the distinction they require. At this point, they tire of the argument, and fall back to a previous formulation of the concept map.

The examples, counter-examples and arguments used by the children were therefore mixed in their effectiveness. The children did not always have sufficient knowledge to carry through the issues raised. In these cases, further resources such as information books, if available, might have helped the children to make progress.

The remoteness of the stars and planets make observation adequate to decide issues, such as that in the previous example, difficult. Experimental manipulation is also largely precluded, although simulations might be of some help on occasions. In one of the groups studying this topic, the children did recall observations of the moon in order to establish the relative distances of the objects in the solar system.

Transcript 43: Group 8 (Earth in space; pre-topic)

P1	Moon is er, Earth is near the moon ... and the moon is near the Earth, innit, *** ?
P2	The moon's all the way up the sky
P1	But sometimes you can see it

For the first topic, Habitats, there were no instances of reference to objects or events, or any other evidence relevant to the concept map. Even though the children had undertaken field studies in which they had observed different habitats and the living things adapted to life there, they did not draw on these experiences to illuminate their discussion. By contrast, the third topic, Sound & hearing, featured quite extensive use of such resources. During both pre- and post-topic sessions, a guitar was available in the classroom, and for the second session, there was also a model ear. Both of these were utilized in the discussion at various times by different groups.

The first example, though, shows how experience from outside school could be drawn on as evidence to support a point being made. Here, the discussion revolved around the connection between the ear drum, hearing and vibration.

Transcript 44: Group 10 (Sound and hearing; pre-topic)

P	For instance, this morning, I was, this is true, I was sitting in the back of the car, and my mum had the radio on too loud, and I could feel it vibrating in the back of my seat.

This first-hand experience makes the clear connection between sound and vibration. The group do not pursue this topic further during the pre-topic session, but return to it when compiling the post-topic map having advanced their understanding somewhat (see Transcript 36). Shortly after this discussion of they way in which sound propagates through the air to the ear drum, the children identified the ear drum on the model ear, and then subsequently the cochlea as "that little twirly thing". They traced the progress of the sound further, identifying the nerves along which the signal passes to the brain. Several of the other groups also used the model as a reminder of the location and relationship of the ear drum and cochlea.

Rather more interesting were the occasions on which the children manipulated objects experimentally, in order to further their understanding. In the first example, one of the children suggests a simple test to answer a query one of the others has raised.

Transcript 45: Group 12 (Sound and hearing; pre-topic)

P1	Do guitar strings, like, vibrate?
P2	Yeah
P3	Yeah
P4	Yeah, it vibrates
P2	N, guitar thing, string vibrates
P3	Yeah it does
P2	You could test it, cos there's a guitar over there

This is a straightforward example of the raising of a question that can be answered by generating empirical evidence, a procedure that underlies the methodology of natural science.

The principal voice drawn on from outside the group was that of the teacher (including the researcher, and in one case, an inspector). The teacher could become engaged in the discussion either as a result of observing the discussion, or by invitation from the children. In some instances, the teacher was clearly seen by the children to be a source of advice, or one who could settle matters they were unsure about. In others, the teacher was more active in prompting the group to rethink or clarify what they had done, or to explore further possibilities.

In group 11's post-topic session, the children had called the researcher to approve their map. Following some requests for clarification regarding the map, the researcher went on to ask them:

> I can think of a connection between those two [pitch and vibration], do you know what it is?

The children were unable to think of any such link, but were able to explain how notes of different pitch could be obtained on a guitar by fretting the strings in different positions. Here is how the discussion proceeded.

Transcript 46: Group 11 (Sound and hearing; post-topic)

Re	Yeah, you can change the pitch of the string by, pressing it in different places
Re	But what's actually happening when you're doing that is you're changing the way the string vibrates
P?	So, mmm
Re	If it's vibrating very very quickly, that's when it's making a high note, and if it's vibrating more slowly, that's when it's making a low note ... it might be worth going and having a look some time when you have the chance ... watching the string vibrating
P1	Shall I go and get the guitar?
Re	Right, you'd better check with *** whether she minds you doing that, but if you look, you'll see that they're vibrating more slowly when they're making low notes
.	
.	
Re	Take that one ... watch how, watch how it vibrates, you can see it going backwards and forwards ... yeah?
P2	Can I do a different one?
Re	Yeah, if you press it up here and make a higher note ... can you see if it's vibrating any quicker?
P3	Yeah
P1	This is a high one
Re	What differences can you see?
Re	Can you feel any difference in the vibration, I've got my hand on there and I can feel
P1	Oh yeah!

Here, the influence of the researcher directs the children to refer to objects and events through which a conceptual relationship can be exemplified. The result is that the children have a vivid demonstration of the correspondence between frequency of vibration and the pitch of the resulting note. In this way, the terms used in the map become rooted in physical actions in the classroom. This is a further example of the point made earlier: that the emerging map provides a record of the discussion and a pivot of negotiation for the teacher which can help to lay bare what has been discussed, as well as gaps that might be filled.

In group 13, an incident occurred that was in some ways parallel to the preceding one. The children were having difficulty in remembering what the cochlea was. They decide to call on their teacher as a resource.

Transcript 47: Group 13 (Sound and hearing; post-topic)

P1	What part of your ear's the cochlea?
Te	Did you look at the ear?
Te	Go and get the model ear
.	
.	
P2	It's this bit isn't it?
Te	Yes
.	
.	
P3	Does this [cochlea] vibrate?
Te	Well, the vibrations come through ... the, ear drum, and then they're sent through these bones here, at the back, and it's this, that changes the vibrations to, um ... messages
P2	So this is the bit that creates them?
P1	Co, co, cochlea
.	
.	
P4	So we could write, the cochlea, um, chan, chan here the, changes, them, to sounds ... changes vibrations to sounds

Once again, a physical object is referred to, in this case a simulation of an otherwise inaccessible part of the body. This enables the nature of the cochlea to be clarified, and also allows the teacher to advance the children's understanding of the way that the different parts of the ear function to create the signals that we interpret as sounds from vibrations in the air.

The use of these resources was intermittent. Sometimes, the resources available enabled new meanings to be negotiated. At other times, the resources were not sufficient, and progress was not made. At their best, the processes at work resembled those in the scientific community, particularly manipulating the physical world to gain empirical data, citing evidence and referring to authorities in the field. It may reasonably be assumed that in these cases, the concept mapping activity played a useful rôle in integrating practical and theoretical perspectives. But such instances were exceptional, and most examples are presented above. This level of integration was not achieved across all the groups. This is an important point to consider in relation to the use of concept mapping to support children's learning in science, and one that will be discussed further in the next chapter.

Summary and conclusions

The children drew on their own prior knowledge, on contacts with the physical world and on the teachers as resources to advance the discussion. At times, a bridge was created in the discussion between the theoretical aspects of the topic being studied and evidence for those theoretical notions. On very rare occasions, experimental manipulation was resorted to in order to generate new evidence. However, there was no systematic attempt on the part of the children to marshal evidence in support of the connections they were formulating. In particular, the epistemic question "what is the evidence for this?" did not arise.

Teachers made use of the window onto the children's discussion afforded by the emerging concept maps to appropriate the discussion so as to advance the children's understanding still further. However, these resources were in relatively short supply, and there were occasions when lack of appropriate resources hindered the development of scientific meaning in the groups. Drawing together various points made in the course of the present chapter, a concern emerges that the concept maps could sometimes become ends in themselves, and their relationship to other work in science could be neglected. The question of how to capitalize on the more successful features of the activity is taken up in the next chapter.

10

DISCUSSION

Since models are a form of conversation, their appreciation in the classroom requires conversation; talking in an active way is indispensable. Clive Sutton (1995, p.6)

10.1 Overview of the Discussion

This thesis is intended to achieve two things. Firstly, it is intended to contribute to an emerging theoretical perspective on the learning of science. Secondly, it is intended to examine how one type of activity, about which there is already a large body of research literature, might support learning construed according to this perspective. The motivation for this research comes from an acknowledged problem in children's coming to apply the constructs and theories of science.

The adoption here of a Wittgensteinian framework to elucidate the nature of this problem is timely. Science educators have begun to point out that the previously dominant "constructivist" ways of characterizing both learning and science have become over-used, masking important questions and insights into learning (O'Loughlin, 1992; Solomon, 1994; Sutton, 1995). One over-emphasized and constricting element has been the individualistic "every man [sic] his own scientist" view of learning (Solomon, op cit. p.7). Underlying this present thesis is the view that there is not, and neither could there be, one true theory of learning in science. Hence the desire to develop what seems to be one promising approach, based on the metaphor of learning a new set of language-games, the language-games of science. This framework liberates research into concept mapping from its narrower roots in Ausubelian theory. It opens up new avenues of exploration, into which the empirical study reported here is a preliminary step. The study takes up first the more familiar question "Does concept mapping have an effect?", and then moves the focus on to the new question "*How* does concept mapping have an effect?".

In 10.2, the main findings of this research will be discussed, and their significance considered in relation to previous findings. This will enable an appraisal of what concept mapping has to offer in the learning of primary science, and the identification of areas where its contribution could be enhanced. Following that, implications for classroom practice will be examined, and suggestions for further development of the approach will be made. 10.3 will reflect on the methodology developed and utilized, leading again to suggestions for further research and development. In 10.4, the broader theoretical position within which this study is located will be brought into focus. Finally, in 10.5, questions for further research are drawn out.

10.2 Research into Concept Mapping

10.2.1 The Main Findings

The various strands of this research, when woven together, lead to four principal conclusions. These will be summarized immediately below, and then expanded upon and discussed in the remainder of the section.

- **Collaborative concept mapping promotes a distinctive form of discourse, which engages children in sustained discussion of scientific ideas and is conducive to learning.**
- **Compared with individual concept mapping, collaborative work helps children focus on scientifically more relevant ideas and produces scientifically superior concept maps.**
- **Ideas that are negotiated collaboratively are more likely to be taken forward by the groups and more likely to be scientifically valid than those generated and developed individually.**

- **There is a danger that concept mapping could become a decontextualized activity, and that links to the wider context of learning in science could be under-utilized.**

Products

In the course of this research into concept mapping, process, product and the relationship between them were investigated. The *product* of one of the sessions observed was a completed concept map for each pupil or group. Within the theoretical position underpinning this study, the concept map is regarded as an inscription that reveals the mappers' grasp of scientific language. Thus the product has a property that is of some interest in addressing the problem alluded to in 10.1: *quality of scientific language*. Each concept map contains some combination of correct conceptions, misconceptions, and everyday or vague connections. The results of the research demonstrate that, as part of an activity carried out after learning about a scientific topic, the concept maps produced by pupils working in collaborative groups were of higher quality than those produced by pupils working individually. They exhibited greater numbers of scientifically appropriate links between constructs (Figures 9.6 to 9.9), and fewer vague associations. This finding is important. Frequently, claims have been made for the advantages of working collaboratively on concept maps. However, previous research has failed to produce any evidence capable of substantiating this point. The findings reported here constitute a first step in assembling relevant evidence by showing collaboration apparently to influence the quality of scientific language transcribed onto the concept map. Within a situation in which they needed to agree on which links to make and how to word them, the collaborative groups seemed to focus on scientifically more relevant connections in the domain.

Discourse structure

However, it is the *process* by which this effect was produced that is of most interest, and which formed the main focus of the research. The underpinning theory locates development of meaning within social interaction. Hence it was hypothesized that features of the linguistic interaction taking place within collaborative groups would account for some of the positive effects noted. An analysis of the speech activities taking place in the groups was therefore developed, to enable the children's talk to be categorised (Chapter 8). The category system proved successful in representing the data as consisting of several different kinds of *discourse move*. On examining how these categories were sequenced and linked into larger structures, three distinctive patterns of *ideational exchange* were identified. What emerged from this analysis was a picture of a style of discourse that differed radically from much of the talk normally characteristic of classroom discussion. The significance of this finding will now be explored.

First, it is worth considering what the discourse was *not*. Although the teacher intervened at times in the dialogue, this occurred infrequently. The talk was largely directed by the children themselves. Hence it avoided what seems to be a universal default of classroom exchanges: the familiar teacher question/student answer/teacher evaluation sequence of "triadic dialogue" (Lemke, 1990). In triadic dialogue, the "question" move does not pose a genuine question at all. It is a prompt for a pupil to produce a response within predetermined limits. Similar dialogue structures occur in all kinds of pedagogical settings, starting from children's very earliest experiences of schooling (Willes, 1983). Reflected in this structure is a tension between, on the one hand, involving pupils in discovering new knowledge for themselves and, on the other, the need to control what it is they discover; hence the initial question may be accompanied by hints as to the required answer (Edwards & Mercer, 1987). For all its pretence of handing over the act of discovery to the pupil, the typical teaching exchange is a highly controlled attempt to bring about a predetermined outcome; imparting knowledge of what Duschl (1994) refers to as "final form science". The main element of discovery involved is directed towards finding out what answer the teacher already has in mind. In contrast, the recordings made during this present study show that it was predominantly *the children* who asked the questions of each other, and that the answers were *not* presupposed (recall that the evaluative moves were of an entirely different sort; see 9.6). Furthermore, the participants often supplied information and offered ideas without any need to be asked, reflecting a very different power structure within the groups to that underlying teacher-pupil interaction.

There was ample evidence that, normally, several speaking turns were involved in introducing an idea and considering it further, often with other members of the group taking a rôle. Figure 9.3 shows that, for nearly all the groups, some quite lengthy exchanges took place as the maps were compiled. These concept mapping sessions, it may therefore be concluded, involved the children in sustained discourse amongst themselves about ideas in science. This in itself is a most significant finding. Lemke has argued for the importance of such sustained discourse in improving learners' ability to use the language of science:

> Students must be given opportunities to speak at greater length (in monologue and dialogue), and to write more, about science topics. The single greatest obstacle to this at present is the dominance of Triadic Dialogue. (*op cit.*, p.168)

The value of talking science

Bennett (1987) points out that much pupil talk in primary classrooms, even during work in groups, does not feature sustained conversation or talk intended to enhance the tasks set. On the evidence reported here, concept mapping provides a welcome contrast. Instead of dialogue in which the admissible answers are fixed in advance, there exists a genuine openness to possible alternatives, with consensus as the arbiter rather than fiat. Such talk may be a vehicle of personal and social discovery in a way that is implicitly denied in triadic dialogue.

O'Loughlin (*op cit.*) argues powerfully, from his reading of Wertsch, that the link between the language environment created in school and the messages given to pupils about their ability to come to know and to act in their own right demands a reappraisal of the rôle of talk in the classroom, of pupils' participation in it and of the power structure underlying it. The kind of discourse promoted by the concept mapping activities examined here would appear to be one way forward in legitimizing and encouraging children's own voices in creating scientific meanings in the classroom.

However, as O'Loughlin acknowledges, there remains a tension between validating pupils' ways of knowing and the responsibility for ensuring they gain access to scientific ways of talking. Discussion is of no value if that which is discussed is of no value. In the context of this research, we therefore have to ask what the discourse was achieving in the way of promoting scientific language. The evidence collected leads to three main conclusions:

- that there was a substantial basis of scientific meanings underlying the talk within the groups;

- that these scientific meanings were maintained and developed through the critical sharing of ideas amongst the members of the groups;
- that where pupils shared responsibility for developing the propositions inscribed in the concept map, these propositions were more likely to be valid scientifically.

The concept mapping activities studied took place as part of programmes of teaching and learning in particular science domains. There were specific links, by means of the terms given to the children, to the scientific ideas covered in those programmes. A major part of the task set for the children was to construct (meaningful) links between those terms in a concept map, and Figure 9.9 shows that the vast majority of the links the children made were based on scientifically acceptable meanings. So it *was* scientific language that the children were engaging with in their discussion. And it is significant that the groups were *discussing* the meaning of the links, rather than taking them as given. Successful discussion depends on communication, and this in turn entails establishing shared meaning. In a collaborative mapping task, the ultimate indication of what is communicable, and therefore taken as shared within the group, is whatever is agreed upon as worth adding to the map. The emerging concept map serves as a means of "semiotic mediation" as Roth & Roychoudhury (1993) put it. Initially, the "map" has little form: only the loose terms that are to be included. But this is enough to establish communication. With the groups studied, much of the task was characterized by the children's considering a range of possible connections between these loose terms. Once a map has begun to take form, the propositions already in place, and their arrangement on the page, serve as a record of what has been agreed so far. However, these children had been encouraged not to fix the form of the maps too early, thereby allowing them to return to links they had already examined, and to make adjustments if necessary. On notable, albeit rare, occasions, the inscription also allowed the teacher a window into the meanings being applied by the children in the group, allowing retrospective access to, and sometimes renegotiation of, connections that had already been discussed (see, for example, Transcript 26). The exercise therefore combines elements of tentative, provisional talk, but also more formalized public outcomes in the form of the finished map. Overall, it can be seen that concept mapping embraces several of the important aspects of learning through communicating identified by Barnes (1976): the feeling of competence gained from validating pupils' own contributions; materials to refer to as common ground for the discussion; an "exploratory" phase; and a public "final draft" towards which the activity proceeds.

At the outset of this research, the kind of talk that upper primary school children would engage in when producing concept maps was not known. Especially, it was not known whether they would take joint responsibility for developing the links, thereby providing opportunities to test whether assumed meanings were actually shared, and to adjust meanings to achieve consensus. This is at the crux of what is meant by "negotiation of meaning" (Barnes, 1976). The data collected for this research included many examples of such negotiative processes.

Figure 9.4 shows how, during a large proportion of the exchanges, the children collaborated in their discussion of a potential link in the concept map. In these cases, not only were children other than the original proposer involved in discussing a link, they were also active in developing its meaning. A commonly recurring pattern, in what were termed "type (iii) exchanges", was for one member of the group to propose that two terms be linked in the map without specifying the meaning of the link, and then for one or more others in the group to elaborate on the connection by stating its meaning (see Transcripts 22 to 24). This is, at first blush, an odd phenomenon. If the originator has spotted a link, why not let her or him explain it? The interpretation offered in 9.5.3 was that this was a way of hypothesis testing to check for shared understanding. The degree of indeterminacy in the way the link was introduced allowed participants to predict the meaning the originator had in mind, and then to confirm this. At the same time, indeterminacy allowed freedom for small adjustments in meaning to be made, and on occasions there was a definite move from an everyday use of terminology to a more scientific usage before the meaning of a link was fixed (9.7). It is in terms very like these that Newman, Griffin & Cole (1989) describe learning in the so-called "zone of proximal development", where knowledge is restructured (see 3.2.4). The implication of this is that the activity went beyond firmly established understanding on the part of the individuals involved, and was therefore engaged in by the pupils at an appropriate level for them to learn from it (Vygotsky, 1978).

Type (iii) exchanges were common, and therefore a highly significant feature of the mode of discourse within the groups (Figure 9.10).

Furthermore, as the figure shows, they were more likely overall to result in incorporation of a scientifically valid proposition in a concept map. In contrast, ideas that were introduced and not elaborated further (as with type (i) exchanges) were considerably less likely to be adopted. Only in one of the domains studied, Earth in space, were these differences not statistically significant. This discrepant result was attributed, in 9.5.1, to the lower level of demand provided by the combination of construct terms the children were given to work with for this topic. For Earth in space, the terms mainly denoted entities of the same ontological category, and could therefore be linked in few ways. In the other topics, there was a greater variety of types of relationship involved, and these seem to have prompted greater collaboration in deciding which connections to make, and elaborating on the nature of these connections. There are therefore indications that aspects of the way the task is set up can affect the kind of discussion that takes place. The implications for classroom practice are discussed in the next section.

However, the implication is, overall, that in the course of concept mapping activities, group processes could be effective in maintaining the discussion around scientific meanings. The maps seem to have provided a structure to the activity supportive of the kinds of language use required for the approach to learning developed in Chapter 3. It is therefore appropriate to turn attention towards the learning gains attributable to the concept mapping engaged in by the children.

Learning gains for individual pupils were the subject of phase two of the research. Learning was evaluated by examining the degree of integration in cognitive structure of concepts pertinent to the scientific domain studied, using Word Association Tests. After studying the topic, the concept mapping group had made significantly greater gains in the number of connections they made between these key ideas than the control group (Table 6.2). This supports the hypothesis that concept mapping contributed to learning for these pupils. However, due to the factors discussed in 6.2.1, there are rival hypotheses capable of explaining this result that cannot be discounted. The results of this phase must therefore be viewed, with due caution, in relation to the large body of previous research that shows the positive contribution made by concept mapping (Chapter 4). That the finding reported here is in keeping with previous results lends strength to the view that concept mapping supported learning in this case too.

Examining in further depth the nature of the links made by pupils in their Word Association Tests indicated that these additional links were interpretable in terms of meaningful learning within the domain studied, and also that there was a definite relationship between the links made in the children's concept maps and those produced later in response to the Word Association Test (6.2.2). Hence the conclusion that the concept mapping activity was partly responsible for the additional gains made by the experimental group gains further credence.

The findings reported here tend to confirm those of previous research that show concept mapping to be an effective learning tool, but also go beyond previous research by providing an explanation of how it achieves the oft-reported learning gains. The explanation offered was that the communicative context of the activity focused attention on appropriately scientific language, which was jointly negotiated. Although this research was into collaborative concept mapping, it may be that, when it is undertaken as an individual activity, some processes are at work in the thinking of the individual equivalent to those identified taking place within the groups. However, the findings reported here show such processes to be less effective in individual than in collaborative work.

The science in the discourse

Through much of the above, it is almost irrelevant that it was *science*, as opposed to other areas of the curriculum, that was being learned. Now it is appropriate to focus more closely on how the processes involved in concept mapping relate to how learning in science was characterized in the early chapters of this thesis.

Science, it will be recalled, was portrayed as consisting in the development of an increasingly powerful descriptive and explanatory language, leading to the discovery of new facts. This language achieves its power through superordinate constructs, and consequently a hierarchical structure to its theories. Some acknowledgement of this is made in the Key Stage 2 programme of study for National Curriculum science, where "Pupils should be given opportunities to ... recognise that science provides explanations for many phenomena" (GB. DFE, 1995, p.7). Underpinning the development of this explanatory language, there is commitment to a set of procedures and standards to justify (though not prove) knowledge claims by means of empirical evidence. Again, there is some recognition of this in the National Curriculum, where there is, in the programme of study for experimental and investigative science, an emphasis on planning for, obtaining and considering evidence (GB. DFE, *op cit.*). Hence the adequacy of the systems of constructs developed by science is gauged by (Putnam, 1981): their comprehensiveness; their economy; their coherence; and the match between what is theoretically predicted and empirical data from observations of the physical universe, particularly (Hacking, 1983) those involving experimental manipulations. In essence, both plausibility and empirical success are required, with theory development leading to testing out of hypotheses, followed again by theory development. It was also noted that the criteria of adequacy of a theory are applied intersubjectively, by the scientific community.

Concept maps, as representations of portions of the theory system, should display coherence and an appropriately hierarchical structure. The evidence from this study confirms that the children were quite capable of producing appropriately structured maps. Not all the domains featured an obviously correct hierarchy, and, as reported in 9.7.1, there were times when the finished map revealed a limited grasp of the theoretical structure of the domain. Yet there was also evidence that the children did, in general, give due consideration to which construct terms should be placed in superordinate positions, showing that they were at least able to put into operation the notion of a hierarchy of constructs. More positively, the finished maps, particularly those made after the topic had been taught, featured coherent networks of relationships. The poorly interrelated chains of construct terms, criticized by Novak & Gowin (1984) as showing a poor grasp of the subject matter, were not a feature of the maps produced during the post-topic sessions. Tentatively, then, we may conclude that these concept maps to some extent reflected, and consequently made explicit, the coherent and interrelated structure characteristic of scientific theories. If the more explicit and coherent set of beliefs is also the more plausible, then concept maps may contribute to the plausibility of scientific theories for the pupils. And as the earlier discussion showed, whether a proposition was approved was a group decision: the criteria of acceptability were being applied intersubjectively.

It is when attention is turned to the relationship between systems of constructs and the collection of empirical data that the position becomes less clear. In a few cases, discussion of the construct terms led to questions' being raised that required reference to physical objects, even to a brief experiment (9.8). These suggested successful integration of the discussion of theory into the whole teaching-learning process. However, whilst pointing up what is possible, such instances were rare, and confined to just one of the topic areas. Much more often, there was no reference made to experience of the physical world. The question "What evidence do we have for this?" did not arise in the course of the discussion.

As structured for these groups, the concept mapping activities clearly did not normally encourage explicit linking of theory to empirical evidence. In a similar way, although there were occasions when the children called on the teacher as a resource to help them resolve a difficulty, this was not common. Neither was there usually any reference made to other secondary sources of information or evidence, such as books. As a result, the discussion occasionally foundered through lack of the appropriate resources to pursue it further.

On the question of suitability for learning science *in particular*, the evidence is therefore mixed. Although the activity seems capable of helping children develop a view of science as a coherent system of constructs with broad explanatory power, this advantage may to some extent have been negated in the cases studied by the tendency for the concept mapping to be isolated from other learning experiences and from the epistemological basis for scientific knowledge. There was evidence that the activity could *potentially* be integrated with empirical work, but this potential was not generally realized. It is not clear how much this tendency to decontextualization was due to the concept mapping's being seen mainly as an assessment tool in the classrooms in which it took place. Certainly, the children were told not to refer to source books, but this does not in itself suffice to isolate the task from other work in science. This therefore raises implications for both classroom practice and further research: how can the activity be enhanced so as to reflect better the processes of science, as well as its products?

10.2.2 Implications for Practice

The research reported in this thesis furthers our understanding of concept mapping as a learning tool, and also enables some recommendations and suggestions to be made as to how the technique's effectiveness may be improved. The evidence presented and discussed above points to one clear implication: concept mapping is an activity that has a positive contribution to make in the learning of science in the latter stages of Key Stage 2. To make best use of it, it is worth considering where the strengths, weaknesses, opportunities and threats lie.

Concept mapping brings, through the parameters of the activity, sufficient structure to discussion. In short, the activity is internally regulated. As such, it can provide a framework within which genuine exploratory discussion can take place and through which children can develop competence in handling scientific language. Part of this structuring is provided by orientation towards a definite outcome, in the form of the final concept map. Furthermore, there is evidence that the discussion converges on an advance in understanding through a meeting of viewpoints. These are strengths, and should commend the activity.

One weakness was the apparent sensitivity of the discussion to the terms used for the mapping. The terms used for the topic Earth in space did not open up such a wide range of kinds of relationship as did the other topics, as discussed above. This was a weakness to the extent that the primary purpose was to generate discussion *for learning*, and it should be recalled that the teacher's purpose in using the concept maps was mainly to obtain assessment information on the children's pre- and post-conceptions. The implication for practice would therefore be to ensure that the range of kinds of construct term chosen was suited to the aims of the activity: terms that can be linked in a variety of different ways would be more likely to generate the kinds of exchange associated here with progress towards scientific meanings. As well as classes of physical object (exemplified here by such terms as "moon" and "guitar string") and of direct experience ("echo" and "vibration"), teachers may consider adding terms denoting entities not observable without instrumentation ("sound wave"), and those without specific referents or with which only pseudo-reference is possible ("survival" and "competition"). It is in elaborating the relationships between such ontological categories that scientific meaning-making progresses (c.f. Harré, 1986), and it is well to reflect this in how we approach the domains to be mapped. As the evidence from this study shows, discussion that focuses on concrete, observable entities can be quite undemanding, and featuring only this type of entity among the terms to be mapped could be expected to encourage trivial links. When the focus is on constructs more removed from concrete experience, one of the things that must be considered in the course of the discussion is how these relate to the observable entities in the map. This should encourage more scientific forms of discussion, in which the relationship between theoretical constructs and evidence features. A first principle for effective discussion of scientific meaning is therefore:

- **Choose terms that reflect the full range of ontological categories in the domain being studied, and not only directly observable entities.**

There was at least one case where an everyday-language term had been given to the children in place of a term with a more specific scientific meaning ("loudness" instead of "volume"). The children seemed to know the term "volume", and to have used loudness in a slightly different way, to refer to *high* volumes. This is a case where progress towards scientific use of language was not served by using non-technical terms, and this too is a point to be considered when choosing words to be mapped. However, the choice of constructs is a decision that must be related to the wider aims in teaching a particular scientific topic, for which the concept mapping is just one means of support.

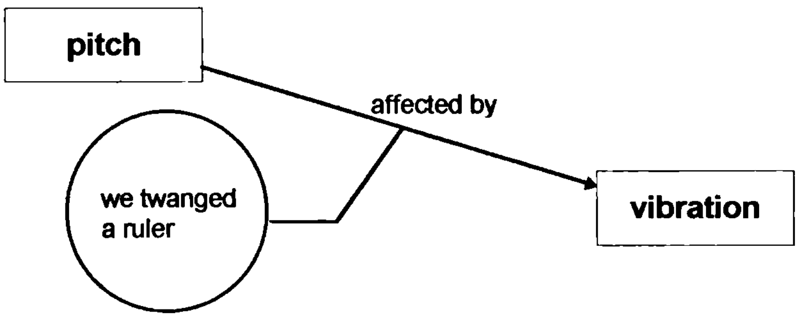
One threat identified was the apparent tendency towards isolation of the concept mapping activity from other work in science. Only rarely did the children refer to previous experiences or to potential investigations. Neither did they seem overly concerned with the evidence for the links they were making. This decontextualization might reduce the overall contribution that the discussion generated could make towards integrated understanding of science. It is therefore necessary to consider ways to counter this weakness. The principle involved is:

- **Seek ways to maintain links between theoretical relationships in the concept maps and evidence for those relationships.**

The following suggestions must remain as speculative responses to this principle until further research has investigated how they function. Questions for further research are raised at the end of the chapter.

An initial possibility would be to ensure that some of the terms given to the children to use in their maps denoted phenomena that counted as evidence for intended relationships in the map. It is likely that children would need some explicit training to work with these, highlighting the purpose of identifying terms denoting evidence and helping them to distinguish between these and terms denoting theoretical terms. A probable difficulty in implementing this approach would be that in many domains appropriate terms would be hard to come up with. What are needed are, effectively, observation statements and references to particular experiences or events, rather than simple labels. Which simple term, for example, would count as denoting evidence for "the moon orbits the Earth"? Consequently, a better approach would probably be to have the children supply their own links to evidence, and this would also help to provoke discussion about the significance of the evidence available. To facilitate this, one could envisage adopting a formalism where, for example, a specific shape of box was used to contain a statement about evidence that was then linked in the concept map to the relationship concerned. Drawing on ideas from the research reported above, this might appear as in Figure 10.1.

Figure 10.1: Example concept-evidence map



This approach would combine some of the elements of Gowin's epistemological "Vee" diagram (Novak & Gowin, 1984) with the concept map. Realizing such an approach in practice may not be straightforward. Kuhn, Amsel & O'Loughlin (1988) concluded that the children they worked with did not always differentiate clearly between theory and evidence. However, perhaps directing children to focus specifically on each can help them make progress in drawing this distinction. Used judiciously, so as not to detract from the overall flexibility of concept mapping, such a modification might enhance the scientific value of the activity.

In this research, concept maps were produced both before and after learning about a scientific topic. Both occasions generated productive discussion (with the exception of the post-topic map for Earth in space, for reasons that have already been mentioned). However, it is possible that this way of scheduling the sessions might itself encourage decontextualization, and that concept mapping could be better integrated with the wider context of work in science by phasing the map-making to be more in step with the phasing of learning. One way would be to use a map as an ongoing working document, to which links were added as particular parts of the domain were explored and the children felt ready to make the links. As an alternative to maps produced by smaller groups, this could be a negotiated product of the whole class, and permanently displayed. It could be combined with the previous suggestion. Using such an approach, the teacher would invite the children to add any links they felt able to, following learning experiences in which some of the constructs were applicable. This would emphasize the relevance of past experiences. With carefully selected construct terms (which could themselves be negotiated with the children) a more immediate connection could be established between the theoretical map and the practical context of objects and events. Novak & Gowin (1984) suggest adding pictures to concept maps to illustrate specific objects and events. Getting children to choose such illustrations could help them to forge links between the theoretical and the practical, between propositions and evidence for their validity.

To bring in, and hence integrate, information from other sources, the children could be encouraged to seek out answers to their queries in, for example, reference material available in the classroom. They could also be encouraged to make use of the teacher's potential to contribute to the discussion, as some groups did in this research. This would be in keeping with the need to value the children's own contributions and attempts to construct meaning, provided they saw the teacher as a resource, as one voice in a conversation, and not as there to control the discussion.

One of the teachers involved in the research made use of plenary sessions in which children from the different groups reported back to the class and they discussed the maps they had made. This is a practice that could help to emphasize the need to aim for a "final draft" as an outcome from the discussion, and consequently the need to adopt publicly defensible scientific language. Another possibility that such sessions could open up would be for the teacher to model the process of making connections between the relationships in the concept map and the wider realm of first-hand experience. Explicit reference could be made to past shared experiences (observations made and investigations conducted). Also, unresolved questions could be identified, either for further investigation or for further information to be located in reference material. The potential for such plenary discussions to contribute to learning is one area where further study is needed, so as to identify the processes at work. If the map were ongoing, they could even help structure a programme of further investigations.

However versatile concept mapping might be, it cannot and should not be the only means of supporting discussion in the classroom. Other activities are needed that provoke the kinds of discourse that concept mapping, at its most effective, provides. In particular, activities are required that prompt children to consider the relationships between theoretical constructions, their application, and evidence. One starting point could be the use of the Proposition Generation Task (Chapter 6) as a focus for group discussion, which in some ways should operate quite similarly to concept mapping. Many other variations on the theme of linking together key terms and describing the relationship are possible. Gilbert's and Pope's (1986) discussion tasks based on cards showing instances and non-instances of a scientific construct provide a different perspective on the use of scientific language.

To focus more on the link between theory and evidence, statements of the relationship between constructs could be what are given to the children, who are then asked to decide on relevant evidence for these. If false links formed part of the list of relationships to be discussed, then the children could supply refuting evidence. Osborne (personal communication, 1995) has proposed a structuring device for discussion of instances, in which pupils are asked to record a classification of given statements into "agree/disagree/don't know", and then to summarize the evidence supporting their classification.

The above suggestions are highly speculative, and leave open many questions. However, it has not been the intention to specify *the* way to conduct concept mapping or discussion tasks. That would be at odds with the thinking behind this research. Rather, it has been the intention to illustrate the various ways in which these activities could be adapted to a range of classroom requirements, and to show how the *principles* of learning science developed in the first part of this thesis may be realized.

10.3 Reflection upon Methodology

The need for more research that goes beyond process-product approaches and looks directly at the way pupils actually tackle tasks in the social environment of their classrooms is acknowledged by Bennett (1987). A major and distinctive feature of this present study has been the development of appropriate methods for investigating the processes involved in group discussion. The methods adopted were of several different kinds, both quantitative and qualitative. However, for this research the methods interlocked, so as to form an overall picture of the nature of the task the children were engaging in. Briefly, the main approaches were:

- use of Word Association Tests in an experimental design to examine changes in children's use of scientific language due to concept mapping;
- scoring concept maps in an experimental design to compare their use as an individual or a group task;
- qualitative analysis of verbal data to characterize the discussion taking place within collaborative concept mapping groups;
- relating identifiable types of talk to the quality of scientific language incorporated in the concept maps produced.

The linchpin of these analyses was the approach to classifying verbal data recorded during pupil discussions. The category scheme that resulted (Figure 8.1) was developed through a process of iteration between established theoretical principles and the raw data gathered. This was successful in categorizing the vast majority of the data (9.3.1). Applying the classification also led to the creation of new constructs to describe the data: the three different types of "ideational exchange". Hence, not only was the description apparently adequate, it also was capable of supporting the development of theory by bringing into view new kinds of entity.

The classification is both theoretically based and grounded in empirical data. The consequence of this is that, while the category system was able to accommodate the range of talk specific to this one activity type, nevertheless these data may be related to other activity types and to a wider range of research on language use. It is likely that this category scheme could be more widely applicable in investigating other types of classroom task, particularly those involving pupil discussion. The approach of "listening in" on pupil talk enabled the identification of relevant processes in the discussion that go some way towards explaining how collaborative concept mapping can support learning, and could well prove fruitful in studying similar discussion-based activities. With further development of the category system to elaborate the different categories of what was termed "contextual talk", the range of uses could be broadened still further. With such refinement, the scheme could provide a general analytic tool supporting classroom research based on verbal data.

In phase two of the study, it was necessary to adopt a means of assessing individual children's cognitive structure in an area of science. The method chosen was the Word Association Test, and this study can be seen as contributing to understanding of this as an assessment approach. Due to logical objections to use of the relatedness coefficient (Garskof & Houston, 1963) to interpret the outcomes in some domains, new methods of analysing the results were developed (6.1.3). These proved particularly valuable in interpreting the Word Association data, providing both a score and a graphical representation of the relationships between the terms used in the test. These methods would be suitable for wider use, including teachers' classroom assessment.

Various pieces of evidence were gathered that indicated use of the Word Association Test to make inferences about the children's understanding in the domain was justified (that is, there is evidence for validity). Some of this evidence accrued from comparisons with other measures: the specially developed Proposition Generation Task and the pupils' concept maps. Other evidence came from qualitative analysis of the lists of responses made by children to the construct terms in the test and the patterns of connections made between them (6.2.1). It was concluded that, not only was the Word Association Test capable of yielding valid inferences, it could also form part of a battery of assessment approaches including the Proposition Generation Task and concept maps. These devices were shown to be addressing closely related, though slightly different, aspects of cognitive organization.

It was also in phase two that some practical difficulties were encountered in implementing the research design as intended. Lack of comparability between the two groups and removal of the opportunity to carry out a second experimental run with the groups reversed made the conclusions from this part of the study more tentative than had originally been anticipated. The threat from such difficulties could have been reduced by opting for a true experiment, with random assignment within the same class, as in the third phase of the research. However, this advantage had in practice to be balanced against the disadvantage of disruption to classroom routine. Disruption would be greater in the case of comparing a concept mapping group to a control group than it would for groups using two different concept mapping approaches. The quest for ideal conditions for an educational experiment always presents a paradox. Life in the primary school is a patchwork of events of different kinds, some predictable and others not: assemblies to rehearse, music tuition for individual pupils, absences for various reasons and serendipitous occurrences that are capitalized upon. Classically, these are sources of "error". But a classroom devoid of such features is not representative of any real context. This research has demonstrated how using a range of methodologies, both quantitative and qualitative, in conjunction can overcome some of the uncertainties introduced by reliance on any specific approach with its inherent weaknesses.

With hindsight, it is possible to suggest one missed opportunity. For two of the classes involved, a plenary session took place after the concept maps had been made, but this was not recorded. In the preceding part of the chapter, the plenary discussion was identified as potentially having an important rôle to play. Had this been known at the time, it might have been possible to record these sessions and thereby to make a start in investigating the extent to which they facilitated links between the concept maps and other work in science. As it is, this must be left to future research.

10.4 Reflection upon Background Theory

The quotation at the beginning of this chapter focuses attention on scientific theorizing as the development of new and more productive ways of talking about the world, which are closely bound up with how we come to see the world. The theoretical basis for this thesis makes a strong connection between human purposes, activities, language and knowledge. As Sutton goes on to say in the same paper:

> language is not just incidental, an after-the-event tool for labelling what we have found, so we can tell someone else about it. Rather, language is interpretive in function, sense-making, theory-constitutive. (*op cit.*, p.4)

It is acknowledgement of this interdependence of ways of talking and ways of seeing that has driven the attempt in this thesis to refocus research into concept mapping onto its rôle in developing children's use of scientific language. Meanings, and consequently ways of seeing, it was argued, are established in communication. Seeing the learning of science in this way involves a Gestalt switch from focusing on the mind of the individual child, striving to make sense of phenomena, to the child as part of a meaning-making community. According priority to social over individual meaning derives from the arguments presented by Wittgenstein (1976). Viewing the individual mind as arising out of "socially rooted and historically developed activities" has its origin in the writings of Vygotsky (1978, p.57).

Development of a scientific viewpoint is (as Sutton points out) unlikely to result from merely informing pupils. Neither, it was argued in Chapter 3, is it likely to result merely from refuting pupils' familiar conceptions. These are strategies for alienation, rather than supporting learning. Instead, it is more likely to arise from negotiating meanings in discussion during which the pupils have some degree of autonomy. Concept mapping was investigated as one way of structuring this kind of discussion, and the results of this have already been considered above. The wider implication of the research reported here is that:

- it is possible for children in Key Stage 2 to collaborate in sustained discussion of scientific meanings, given the right support.

Also, it seems that this discussion can have a positive effect on the children's ability to use scientific language. The different types of ideational exchange identified as characterizing this discourse, and their rôles in developing a shared response to the task, are likely to be evident in other kinds of discussion too. Exchanges in which an idea was developed jointly by several members of a group seemed to be a particularly important feature of the discussion (see 10.2), and can now be seen to be implicated in other successful cooperative learning initiatives. These findings correspond, for example, with those reported by Webb (1982a, 1982b, 1989) that show the sharing of explanations between pupils collaborating on a problem-solving task (in mathematics or computer science) to be beneficial to their learning. King (1990) also reports positive effects from what is referred to here as collaborative elaboration, structured, in the case of her students, by written prompts.

As suggested above, the work reported here also builds on Barnes' early (1976) work, in which he proposed the usefulness of exploratory discussion, working towards a public "final draft" outcome. In his later work with Todd (Barnes & Todd, 1977, 1981), the following questions were addressed:

> Was it possible to describe the talk of small groups of children in ways which would distinguish successful engagement with the learning tasks they had been set? And if that were possible, could one identify aspects of the social context and of the setting of the task which relate to successful learning? In particular, could those moves by which learners interrelate their viewpoints be identified? (1981, p.69)

They encountered considerable difficulty in doing so, and were able to find no simple relationships between the form of what was said (its wording, grammatical description and so on) and the strategies being enacted. This present study has benefited from their (unfinished) work on analysis at the level of discourse moves. It has been possible to take a significant step forward in addressing those questions by identifying sequences of discourse move which seem to be characteristic of learners' interrelating their points of view and of their successful engagement with the task. At present, this has been achieved with one kind of discussion task, which structures talk by providing a restricted range of possibilities (a set of terms to link) and autonomy in discussing how they should be linked. Clearly more work is needed with other tasks before there can be confident answers to these questions, but a positive beginning has been made.

At the same time, concern was expressed regarding the "aboutness" of the dialogue recorded. It was not always convincingly the case that the communication established was unambiguously related to a wider realm of experience and meaning beyond the confines of the concept mapping activity itself. Hence the emphasis placed, in this chapter, on *evidence* and on the integration of learning experiences, for these must now be seen as key to the apprenticeship notion of learning developed in Chapter 3. Lemke (1990) raises a point pertinent to this issue:

> The rhetoric of "evidence and proof" presumes that evidence itself simply exists, is found simply "by looking". It conveniently ignores that *people* always have to *decide* that something will count as evidence for something else. (p.142)

How we come to make decisions about evidence and its relationship to theory is therefore of central concern if *education* in science, rather than learning how to get by in school science, is a serious aim. Personal engagement in coordinating theory and evidence is needed, and the rôle of the pupil's own voice in this process cannot be marginalized. Consequently, Lemke concludes:

> We must teach students that scientific conclusions are always fallible human judgements, not absolute facts, and that science as a whole is a messy, human business, not a perfect method for discovering absolute truth. We are going to have to give students practice at using science, together with an appreciation of differences in social values and interests, to make decisions about real issues. (*ibid.*, p.150)

This cannot be achieved in a classroom environment in which there is one authoritative voice, that of science, often speaking vicariously through the teacher. The children must be enabled to make informed judgements based on evaluation of the evidence. In turn, this requires an appreciation of the epistemology of science, rather than unquestioning acceptance of its products. This study has made a contribution to understanding how the processes involved in group discussion, as manifest in one type of relatively open classroom task, might relate to the wider aims of inducting children into scientific epistemology and language. But there are deeper questions that need to be raised. Given that primary education is conducted under conditions of compulsion, with the teacher in an *ex officio* position of authority, can children make a distinction in practice between the introduction of "useful new ways of talking" to supplement their preconceptions and implicit denial of the validity of those preconceptions? Do they see these new exciting alternatives as just another imposition? The answers to these questions will depend on the kind of autonomy children actually experience in the course of their learning in science. Ultimately, success in this enterprise is dependent on creating a learning environment in which the activities engaged in by pupils (and not only the words) are meaningful to them. Edwards & Mercer (1987) observe that:

> The experiences and activities of the classroom are made meaningful by the sense made of those things by classroom talk. When teachers go out of their way to avoid offering to pupils overt help in making sense of their experiences, the consequences may be that the usefulness of those experiences is lost, or that the teacher and pupils resort to more surreptitious means of communicating what is conventional sense. So we find teachers asking questions and miming the answers. For many pupils, learning from teachers must appear to be a mysterious and arbitrarily difficult process, the solution to which may be to concentrate on trying to do and say what appears to be expected - a basically "ritual" solution. (p.169)

Here, the problem of alienation and ritual can be traced back to over-emphasis on the individual's response, for it is individuals who provide the "right answer" to the teacher's pseudo-questions (insights that could be detected in the writings of Holt, 1964, a quarter-century earlier). But science is a *collaborative* enterprise. Scientific objectivity is dependent on the communicability of evidence and the *process* of intersubjectivity (Solomon, 1994). Thus it is the contention of the present author that adopting a scientific "way of talking" can indeed be other than an imposition, provided we shift the focus from individual sense-making. Operationalizing this view in the classroom demands a move further towards collaborative enquiry. Some initial suggestions have been made as to how discussion-based tasks may empower pupils by helping them *jointly* to gain access to the sorts of cultural tools they need in order to act as critical consumers of scientific knowledge claims. Further research is needed on how increased emphasis on collaborative meaning-making can affect children's perception of science as a way of knowing and of themselves as learners.

10.5 Directions for Further Research

The preceding discussion has touched on several areas in which this research may be furthered. In particular, some new emphases in the use of concept mapping in the classroom have been suggested. Some of the chief questions raised will now be reiterated. In parallel with these suggested improvements, it is also desirable to validate the findings reported in this thesis in a wider range of circumstances, so as to ascertain the limits of their generalizability.

- *Can concept mapping be adapted to help children establish closer links between theoretical interpretations and phenomena?*

Ways to help children focus on evidence and to create stronger links between discussion and other work in science need to be found and evaluated. The use of concept maps as working documents that are updated regularly in order to reflect on experience, and emphasis on plenary discussion are potentially fruitful approaches. The methods used in this study could be used to pursue this question, but the scope of the research could usefully be extended to embrace the pupils' perceptions of the activity and what they felt they gained from it.

Broadening the scope of the enquiry still further, the following question may be raised.

- *What other activity types promote the kind of discourse identified?*

The analysis scheme developed here could be applied to a range of classroom activities, and some suggestions were made in 10.2.2 of tasks that could be researched. The results would have practical implications in identifying specific useful activities that promote learning. Collaborative meaning-making was pinpointed above as possibly being a key element in developing a deep appreciation of the scientific "way of seeing the world". But discussion of scientific ideas is still only a partially understood aspect of learning. Hence how teachers can make best use of discussion tasks becomes of interest, together with how this use impacts on the learning environment in the classroom and how pupils perceive it. An important question is:

- *What effect does a move towards more negotiated development and application of scientific meanings have on the overall learning environment?*

10.6 Concluding Remarks

This work began with reference to a problem: that of children's tendency to hold scientifically inappropriate ideas in spite of teaching. Over subsequent chapters, this problem was elucidated as being one of a lack of *meaning* of the language of science for these pupils. Concept mapping was investigated as a means of accessing that meaning. Eger (1992), whilst considering the problems of learning in a specific domain, provides a convenient summary of the thinking behind this investigation.

> What is involved is a struggle by the newcomer for a clear view of the ontological landscape, in which forces, being relational terms, are different in kind from entities used to define the state - like energies and momenta. Today, this struggle, a recapitulation of the original Newtonian struggle, is made easier by the fact that an appropriate language already exists. But, as many educators now realize, to understand this language it is not enough merely to learn the definitions of the terms, and go over a few examples of their use. That is why there is an increasing demand for "real world" experiences in education, an end to "lecturing", and so on. (p.341)

He goes on to highlight the problem of decontextualization of learning tasks from the "real world" of experience, urging us to find ways to remake the link:

> Despite a whole series of reforms during the past decades, the feeling remains, in science and elsewhere, that for some reason the *study* of things is still remote from the things studied; that it does not *enter into* those things, but deals with them from a distance. Hermeneutics suggests here a failure of interpretation - not of *translation*, which can be passive or automatic, but of interpretation as a mode of *being in* that which is interpreted. (*ibid.* p.341-342)

Concept mapping was found to be an effective way to give children a better grasp of the language of science. With some more attention to how it is used in the classroom, it offers the further possibility of assisting children to enter fully into that language, and moreover, through *using* that language, to enter more fully into the world by seeing it in new ways.

BIBLIOGRAPHY

ABAYOMI, B. I. (1988). The effects of concept mapping and cognitive style on science achievement. *Dissertation Abstracts International, 49*(6), 1420A

ARCHENHOLD, W. F., DRIVER, R. H., ORTON, A., & WOOD-ROBINSON, C. (Eds.) (1980). *Cognitive Development Research in Science and Mathematics.* Leeds: Leeds University

AULT, C. R. (1985). Concept mapping as a study strategy in earth science. *Jnl. of College Science Teaching, 15*(1), 38-44

AUSUBEL DP, NOVAK JD & HANESIAN H (1978). *Educational Psychology: a cognitive view* (2nd edn). New York: Holt, Rinehart, Winston

BAIRD JR & MITCHELL IJ (1987) (Eds.). *Improving the Quality of Teaching and Learning.* Melbourne: Monash University

BANNISTER D & FRANSELLA F (1971). *Inquiring Man.* Harmondsworth: Penguin

BAR-LAVIE BZ (1988). Enhancing meaningful learning in an environmental education program: a case study of a class empowered through the use of Novak's and Gowin's principles of learning how to learn, concept mapping, interviewing and educating. *Dissertation Abstracts International, 48*(10) 2590A

BARENHOLZ H & TAMIR P (1992). A comprehensive use of concept mapping in design instruction and assessment. *Research in Science and Technological Education, 10*(1) p.37-52

BARNES D (1976). *From Communication to Curriculum.* Harmondsworth: Penguin

BARNES D & TODD F (1977). *Communication and Learning in Small Groups.* London: Routledge & Kegan Paul

BARNES D & TODD F (1981). Talk in small learning groups: analysis of strategies. In Adelman C (Ed.). *Uttering, Muttering: collecting, using and reporting talk for social and educational research.* London: Grant McIntyre

BENNETT N (1987). Changing perspectives on teaching-learning processes in the post-Plowden era. *Oxford Review of Education, 13*(1) p.67-79

BENNETT N, DESFORGES C, COCKBURN A & WILKINSON B (1984). *The Quality of Pupil Learning Experiences.* London: Lawrence Erlbaum

BORG WR & GALL MD (1989). *Educational Research: an introduction* (5th edn). New York: Longman

BROWN A (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In Weinert FE & Kluwe RH (Eds.). *Metacognition, Motivation, and Understanding.* Hillsdale: Lawrence Erlbaum

BULMAN L (1985). *Teaching Language and Study Skills in Secondary Science.* London: Heinemann

BUZAN T (1974). *Use Your Head.* London: BBC

CACHAPUZ AFC & MASKILL R (1989). Using word association in formative classroom tests: following the learning of Le Chatelier's principle. *International Jnl. of Science Education, 11*(2) p.235-246

CAREY S (1985). *Conceptual Change in Childhood.* Cambridge: MIT Press

CARMICHAEL P, DRIVER R, HOLDING B, PHILLIPS I, TWIGGER D & WATTS M (1990). *Research on Students' Conceptions in Science: a bibliography.* Leeds: Children's Learning in Science Group, University of Leeds

CHI MTH & KOESKE RD (1983). Network representation of a chilld's dinosaur knowledge. *Developmental Psychology, 19*(1) p.29-39

CLEARE CC (1983). Using concept mapping to detect interventions effective in improving pre-service elementary education majors' understanding of science topics. In Helm H & Novak JD (Eds.). *Proceedings of the International Seminar Misconceptions in Science and Mathematics.* New York: Cornell University

COCHAUD GAM & THOMPSON JJ (1980). Sequencing science material for particular classes. In Archenhold WF, Driver RH, Orton A & Wood-Robinson C (Eds.). *Cognitive Development Research in Science and Mathematics.* Leeds: Leeds University

COHEN D (1987). The use of concept maps to represent unique thought processes: toward more meaningful learning. *Jnl. of Curriculum and Supervision, 2*(3) p.285-289

COHEN J (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1) p.37-46

COHEN L & MANION L (1980). *Research Methods in Education.* London: Croom Helm

COLLINS AM & QUILLIAN MR (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8*(2) p.240-247

COLLINS AM & LOFTUS EF (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6) p.407-428

CONSTAS MA (1992). Qualitative analysis as a public event: the documentation of category development procedures. *American Educational Research Jnl., 29* (2) p.253-266

COSGROVE M & OSBORNE RJ (1985). Lesson frameworks for changing children's ideas. In Osborne RJ & Freyberg PS (Eds.). *Learning in Science: the implications of children's science.* Aukland: Heinemann

CURRY KJW (1986). A study of the value of concept mapping for third grade students. *Masters Abstracts International, 24*(3)

DRIVER R (1983). *The Pupil as Scientist?* Milton Keynes: Open University Press

DRIVER R & EASLEY J (1978). Pupils and paradigms: a review of literature related to concept development in adolescent science students. *Studies in Science Education, 5* p.61-84

DRIVER R & OLDHAM V (1985). A Constructivist Approach to Curriculum Development. *Studies in Science Education, 13* p.105-122

DRIVER R, GUESNE E & TIBERGHIEN A (1985). Some features of children's ideas and their implications for teaching. In Driver R, Guesne E & Tiberghien A (Eds.). *Children's Ideas in Science.* Milton Keynes: Open University

DRIVER R, ASOKO H, LEACH J, MORTIMER E, SCOTT P (1994). Constructing scientific knowledge in the classroom. *Educational Researcher, 23*(7) p.5-12

DUSCHL RA (1994). Research on the history and philosophy of science. In Gabel DL (Ed.). *Handbook of Research on Science Teaching and Learning.* New York: Macmillan

EDWARDS D & MERCER N (1987). *Common Knowledge: the development of understanding in the classroom.* London: Methuen

EGER M (1992). Hermeneutics and science education: an introduction. *Science and Education, 1*(4) p.337-348

EGER M (1993a). Hermeneutics as an approach to science: part I. *Science and Education, 2*(1) p.1-29

EGER M (1993b). Hermeneutics as an approach to science: part II. *Science and Education, 2*(4) p.303-328

ERICKSON BH & NOSANCHUK TA (1979). *Understanding Data.* Milton Keynes: Open University Press

EYLON B & LINN MC (1988). Learning and instruction: an examination of four research perspectives in science education. *Review of Educational Research, 58*(3) p.251-301

EYLON B & RIEF F (1984). Effects of knowledge organization on task performance. *Cognition and Instruction, 1*(1) p.5-44

FISHER KM (1986). *Elaboration of Cognitive Knowledge of Biology from Childhood to Adulthood.* Research Report, University of California

GAGNÉ RM & WHITE RT (1978). Memory structures and learning outcomes. *Review of Educational Research, 48*(2) p.187-222

GARNER R (1990). When children and adults do not use learning strategies: toward a theory of settings. *Review of Educational Research, 60*(4) p.517-529

GARSKOF BE & HOUSTON JP (1963). Measurement of verbal relatedness: an idiographic approach. *Psychological Review, 70*(3) p.277-288

GILBERT JK & POPE ML (1986). Small group discussions about conceptions in science: a case study. *Research in Science and Technological Education, 4*(1) p.61-76

GLASERSFELD E VON (1989). Learning as a Constructive Activity. In Murphy P & Moon B (Eds.). *Developments in Learning and Assessment.* London: Hodder & Stoughton

GLYNN SM, YEANY RH & BRITTON BK (1991). A constructive view of learning science. In Glynn SM, Yeany RH & Britton BK (Eds.). *The Psychology of Learning Science.* Hillsdale: Lawrence Erlbaum

GOETZ JP & LECOMPTE MD (1984). *Ethnography and Qualitative Design in Educational Research.* Orlando: Academic Press

GOLDSMITH TE, Johnson PJ & Acton WH (1991). Assessing structural knowledge. *Jnl. of Educational Psychology, 83*(1) p.88-96

GOODMAN N (1978). *Ways of Worldmaking.* Indianapolis: Hackett

GREAT BRITAIN. DEPARTMENT OF EDUCATION AND SCIENCE & WELSH OFFICE (1991). *Science in the National Curriculum.* London: HMSO.

GREAT BRITAIN. DEPARTMENT FOR EDUCATION & WELSH OFFICE (1995). *Science in the National Curriculum.* London: HMSO.

GUILFORD JP (1981). *Fundamental Statistics in Psychology and Education.* (6th edn) New York: McGraw Hill

GUNSTONE RF (1980). Word association and the description of cognitive structure. *Research in Science Education, 10* p.45-53

HACKING I (1983). *Representing and Intervening.* Cambridge: Cambridge University Press.

HALLIDAY MAK (1973). *Explorations in the Functions of Language.* London: Edward Arnold

HALLIDAY MAK (1978). *Language as Social Semiotic.* London: Edward Arnold

HAMMERSLEY M (1991). *Deconstructing the Qualitative and Quantitative Divide.* Paper presented at the BERA annual conference

HANF MB (1971). Mapping: a technique for translating reading into thinking. *Jnl. of Reading, 14* p.225-230, 270

HARLEN W (1990). *Learning science and doing science.* Proceedings of the Annual Conference of the National Foundation for Educational Research. Slough: NFER

HARRÉ R (1986). *Varieties of Realism.* Oxford: Blackwell

HARRIS FO & FERGUSON HJC (1979). *Chemistry 11-13.* Harlow: Longman

HARRISON B (1979). *An Introduction to the Philosophy of Language.* London: Macmillan

HEINZE-FRY JA & NOVAK JD (1990). Concept mapping brings long-term movement toward meaningful learning. *Science Education, 74*(4) p.461-472

HEWSON MGA'B (1985). The rôle of intellectual environment in the origins of conceptions: an exploratory study. In West LHT & Pines AL (Eds.). *Cognitive Structure and Conceptual Change.* Orlando: Academic Press

HUITEMA BE (1980). *The Analysis of Covariance and Alternatives.* New York: John Wiley & Sons

HOLT J (1964). *How Children Fail.* Harmondsworth: Penguin

HORTON PB, MCCONNEY AA, GALLO M, WOODS AL, SENN GJ & HAMELIN D (1993). An investigation of the effectiveness of concept mapping as an instructional tool. *Science Education, 77*(1) p.95-111

ISA AM & MASKILL R (1982). A comparison of science word meaning in the classrooms of two different countries: Scottish Integrated Science in Scotland and Malaysia. *British Jnl. of Educational Psychology, 52*(2) p.188-198

JACOB E (1987). Qualitative research traditions: a review. *Review of Educational Research, 57*(1) p.1-50

JEGEDE OJ, ALAIYEMOLA FF & OKEBUKOLA PAO (1990). The effect of concept mapping on students' anxiety and achievement in biology. *Jnl. of Research in Science Teaching, 27*(10) p.951-960

JOHNSON JK & HOWE AC (1978). The use of cognitive conflict to promote conservation acquisition. *Jnl. of Research in Science Teaching, 15*(4) p.239-247

JOHNSON-LAIRD PN (1985). Mental models. In Aitkenhead AM & Slack JM (Eds.). *Issues in Cognitive Modeling.* London: Lawrence Erlbaum

KANT I (1993). *Critique of Pure Reason,* translated by Meiklejohn and edited by V Politis. London: Everyman

KEMPA RF & AYOB A (1991). Learning interactions in group work in science. *International Jnl. of Science Education, 13*(3) p.341-354

KEMPA RF (1993). *Science education research: from theory to practice.* Paper presented at the International Conference on Science Education in Developing Countries, Tel Aviv

KERLINGER FN & PEDHAZUR EJ (1973). *Multiple Regression in Behavioral Research.* New York: Holt Rinehart & Winston

KING A (1990). Enhancing peer interaction and learning in the classroom through reciprocal questioning. *American Educational Research Jnl., 27*(4) p.664-687

KING A (1992). Comparison of self questioning, summarizing, and notetaking-review as strategies for learning from lectures. *American Educational Research Jnl., 29*(2) p.303-323

KNORR-CETINA KD (1981). *The Manufacture of Knowledge.* Oxford: Pergamon

KUHN D, AMSEL E & O'LOUGHLIN M (1988). *The Development of Scientific Thinking Skills.* San Diego: Academic Press

KUHN TS (1970a). Logic of discovery or psychology of research? In Lakatos I & Musgrave A (Eds.). *Criticism and the Growth of Knowledge.* Cambridge: Cambridge University Press

KUHN TS (1970b). Reflections on my critics. In Lakatos I & Musgrave A (Eds.). *Criticism and the Growth of Knowledge.* Cambridge: Cambridge University Press

KUHN TS (1970c). *The Structure of Scientific Revolutions.* Chicago: University of Chicago

LAKATOS I (1970). Falsification and the methodology of scientific research programmes. In Lakatos I & Musgrave A (Eds.). *Criticism and the Growth of Knowledge.* Cambridge: Cambridge University Press

LAKOFF G (1987). *Women, Fire, and Dangerous Things: what categories reveal about the mind.* Chicago: University of Chicago

LEHMAN JD, CARTER C & KAHLE JB (1985). Concept mapping, Vee mapping, and achievement: results of a field study with black high school students. *Jnl. of Research in Science Teaching, 22*(7) p.663-673

LEMKE JL (1990). *Talking Science: language, learning and values.* Norwood: Ablex

LIGHT P & GLACHAN M (1985). Facilitation of individual problem solving through peer interaction. *Educational Psychology, 5*(3+4) p.217-225

LINN RL (1988). Change assessment. In Keeves JP (Ed.). *Educational Research, Methodology and Measurement: an international yearbook.* Oxford: Pergamon

LONNING RA (1993). Effect of cooperative learning strategies on student verbal interactions and achievement during conceptual change instruction in 10th grade general science. *Jnl. of Research in Science Teaching, 30*(9) p.1087-1101

MCALEESE R (1985). Some problems of knowledge representation in an authoring environment: exteriorization, anomolous state meta-cognition and self-confrontation. *Programmed Learning and Educational Technology, 22*(4) p.299-306

MALONE J & DEKKERS J (1984). The concept map as an aid to instruction in science and mathematics. *School Science and Mathematics, 84*(3) p.220-232

MATTHEWS MR (1992). *Constructivism and empiricism: an incomplete divorce.* Paper presented at the ASERA conference, Hamilton

MERRILL WL (1987). Preservice teachers' understanding of division as assessed by concept mapping. *Dissertation Abstracts International, 45*(5) 1175A

MESSICK S (1989). Validity. In RL Linn (Ed). *Educational Measurement* (3rd edn). New York: ACE-Macmillan

MILLAR R (1989). Constructive criticisms. *International Jnl. of Science Education, 11* (Special Issue) p.587-596

MILLER GA (1991). *The Science of Words.* New York: Scientific American Library

MOREIRA MA & SANTOS CA (1981). The influence of content organization on students' cognitive structure in thermodynamics. *Jnl. of Research in Science Teaching, 18*(6) p.525-531

MUTHUKRISHNA N, CARNINE D, GROSSEN B & MILLER S (1993). Children's alternative frameworks: should they be directly addressed in science instruction? *Jnl. of Research in Science Teaching, 30*(3) p.233-248

NEWMAN D, GRIFFIN P & COLE M (1989). *The Construction Zone: working for cognitive change in school.* Cambridge: Cambridge University Press

NORUSIS MJ (1988). *SPSS/PC+ Studentware Manual.* Chicago: SPSS Inc

NOVAK JD (1977). An alternative to Piagetian psychology for science and mathematics education. *Science Education, 61*(4) p.453-477

NOVAK JD (1983). Can metalearning and metaknowledge strategies to help students learn how to learn serve as a basis for overcoming misconceptions? in Helm H & Novak JD (Eds.). *Proceedings of the International Seminar Misconceptions in Science and Mathematics.* Cornell University

NOVAK JD (1990a). Concept maps and Vee diagrams; two metacognitive tools to facilitate meaningful learning. *Instructional Science, 19* p.29-52

NOVAK JD (1990b). Concept mapping: a useful tool for science education. *Jnl. of Research in Science Teaching, 27*(10) p.937-949

NOVAK JD & GOWIN DB (1984). *Learning How to Learn.* Cambridge: Cambridge University Press

NOVAK JD, GOWIN DB & JOHANSEN GT (1983). The use of concept mapping and knowledge vee mapping with junior high school science students. *Science Education, 67*(5) p.625-645

NOVAK JD & MUSONDA D (1991). A twelve-year longitudinal study of science concept learning. *American Educational Research Journal, 28*(1) p.117-153

O'LOUGHLIN M (1992). Rethinking science education: beyond Piagetian constructivism toward a sociocultural model of teaching and learning. *Jnl. of Research in Science Teaching, 29*(8) p.791-820

OGBORN J (1995). Recovering reality. *Studies in Science Education,* 25 p.3-38

OKEBUKOLA PA (1990). Attaining meaningful learning of concepts in genetics and ecology: an examination of the potency of the concept mapping technique. *Jnl. of Research in Science Teaching, 27*(5) p.493-504

OKEBUKOLA PA & JEGEDE OJ (1988). Cognitive preference and learning mode as determinants of meaningful learning through concept mapping. *Science Education, 72*(4) p.489-500

OSBORNE J (1993). *Beyond constructivism.* Paper presented at the Third International Seminar on Misconceptions and Educational Strategies in Science and Mathematics, Cornell University

OSBORNE J (1995). Personal communication

OSBORNE RJ & FREYBERG PS (Eds.)(1985). *Learning in Science: the implications of children's science.* Auckland: Heinemann

OSBORNE RJ & WITTROCK MC (1983). Learning science: a generative process. *Science Education, 67*(4) p.489-508

PANKRATIUS WJ (1990). Building an organized knowledge base: concept mapping and achievement in secondary school physics. *Jnl. of Research in Science Teaching, 27*(4) p.315-333

PESHKIN A (1993). The goodness of qualitative research. *Educational Researcher, 22*(2) p.23-29

PFUNDT H & DUIT R (1994). *Bibliography: Students' Alternative Frameworks and Science Education* (4th edn). Kiel: Kiel University/IPN

POPPER KR (1970). Normal science and its dangers. In Lakatos I & Musgrave A (Eds.). *Criticism and the Growth of Knowledge.* Cambridge: Cambridge University Press

PRAWAT RS (1989). Promoting access to knowledge, strategy and disposition in students: a research synthesis. *Review of Educational Research, 59*(1) p.1-41

PREECE PFW (1976). Mapping cognitive structure: a comparison of methods. *Jnl. of Educational Psychology, 68*(1) p.1-8

PRIEST S (1991). *Theories of the Mind.* London: Penguin

PUTNAM H (1981). *Reason, Truth and History.* Cambridge: Cambridge University Press

RORTY R (1980). *Philosophy and the Mirror of Nature.* Oxford: Blackwell

ROTH W-M & ROYCHOUDHURY A (1992). The social construction of scientific concepts or the concept map as conscription device and tool for social thinking in high school science. *Science Education, 76*(5) p.531-557

ROTH W-M & ROYCHOUDHURY A (1993). The concept map as a tool for the collaborative construction of knowledge: a microanalysis of high school physics students. *Jnl. of Research in Science Teaching, 30*(5) p.503-534

ROTH W-M & ROYCHOUDHURY A (1994). Science discourse through collaborative concept mapping: new perspectives for the teacher. *International Jnl. of Science Education, 16*(4) p.437-455

ROWELL JA & DAWSON CJ (1989). Towards an integrated theory and practice for science teaching. *Studies in Science Education, 16* p.47-73

SAINSBURY MJ (1992). *Meaning, Communication and Understanding in the Classroom.* Aldershot: Avebury

SAMARAPUNGAVAN A (1991). *Children's metajudgements in theory choice tasks: scientific rationality in childhood.* Paper presented at the annual conference of the American Educational Research Association, Chicago

SCHMID R & TELARO G (1990). Concept mapping as an instructional strategy for high school biology. *Jnl. of Educational Research, 84*(2) p.78-85

SEARLE JR (1969). *Speech Acts: an essay in the philosophy of language.* Cambridge: Cambridge University Press

SHAVELSON RJ (1972). Some aspects of the correspondence between content structure and cognitive structure in physics instruction. *Jnl. of Educational Psychology, 63*(3) p.225-234

SHAVELSON RJ (1974). Methods for examining representations of a subject-matter structure in a student's memory. *Jnl. of Research in Science Teaching, 11*(3) p.231-249

SHAVELSON RJ & STANTON GC (1975). Construct validation: methodology and application to three measures of cognitive structure. *Jnl. of Educational Measurement, 12*(2) p.67-85

SHUELL TJ (1986). Cognitive conceptions of learning. *Review of Educational Research, 56*(4) p.411-436

SILVERMAN D (1993). *Interpreting Qualitative Data: methods for analysing talk, text and interaction.* London: Sage

SINCLAIR J McH & COULTHARD RM (1975). *Towards an Analysis of Discourse.* London: Oxford University Press

SMITH F (1971). *Understanding Reading.* London: Holt, Rinehart & Winston

SMITH MU (1992). Expertise and the organization of knowledge: unexpected differences among genetic counsellors, faculty, and students on problem categorization tasks. *Jnl. of Research in Science Teaching, 29*(2) p.179-205

SNIR J (1991). Sink or float- what do the experts think?: The historical development of explanations for flotation. *Science Education, 75*(5) p.595-609

SOLOMON J (1983). Learning about energy: how pupils think in two domains. *European Jnl. of Science Education, 5*(1) p.49-59

SOLOMON J (1994). The rise and fall of constructivism. *Studies in Science Education, 23* p.1-19

STENSVOLD MS & WILSON JT (1990). The interaction of verbal ability with concept mapping in learning from a chemistry laboratory activity. *Science Education, 74*(4) p.473-480

STEWART J (1979). Content and cognitive structure: critique of assessment and representation techniques used by science education researchers. *Science Education, 63*(3) p.395-405

STEWART J (1980). Techniques for assessing and representing information in cognitive structure. *Science Education, 64*(2) p.223-235

STEWART J, VAN KIRK J & ROWELL R (1979). Concept maps: a tool for use in biology teaching. *The American Biology Teacher, 41*(3) p.171-175

STICE CF & ALVAREZ MC (1986). *Hierarchical Concept Mapping: young children learning how to learn (a viable heuristic for the primary grades).* Report No.5 Center for Excellence, Basic Skills for the Disadvantaged, Reading/Writing component. Nashville: Tennessee State University

STICE CF & ALVAREZ MC (1987). Hierarchical concept mapping in the early grades. *Chidhood Education, 64*(2) p.86-96

STRAUSS A & CORBIN J (1990). *Basics of Qualitative Research: grounded theory procedures and techniques.* Newberry Park: Sage

STUBBS M (1981). Scratching the surface. In Adelman C (Ed.). *Uttering, Muttering: Collecting, Using and Reporting Talk for Social and Educational Research.* London: Grant McIntyre

STUBBS M (1983). *Discourse Analysis.* Oxford: Blackwell

SUCHTING WA (1992). Constructivism deconstructed. *Science and Education, 1*(3) p.223-254

SUBKOVIAK MJ (1980). Decision-consistency approaches. In Berk RA (Ed.). *Criterion Referenced Measurement: the state of the art.* Baltimore: John Hopkins University Press

SUTTON CR (1980). The learner's prior knowledge: a critical review of techniques for probing its organization. *European Jnl. of Science Education, 2*(2) p.107-120

SUTTON CR (1992). *Words Science and Learning.* Buckingham: Open University

SUTTON CR (1995). *The scientific model as a form of speech.* Paper presented at the European Conference for Research in Science Education, University of Leeds

SYMINGTON DJ & NOVAK JD (1982). Teaching children how to learn. *Educational Magazine, 39*(5) p.13-16

TRAUB RE (1994). *Reliability for the Social Sciences.* Thousand Oaks: Sage

VEDDER PH (1985). *Cooperative Learning: a study on processes and effects of cooperation between primary school children.* The Hague: SVO

VYGOTSKY LS (1978). *Mind in Society: the development of higher psychological processes,* edited by M Cole *et al.* Cambridge: Harvard University Press

WALKER R & ADELMAN C (1976). Strawberries. In Stubbs M & Delamont S (Eds.). *Explorations in Classroom Observation.* London: John Wiley

WALLACE JD & MINTZES JJ (1990). The concept map as a research tool: exploring conceptual change in biology. *Jnl. of Research in Science Teaching, 27*(10) p.1033-1052

WEBB NM (1982a). Peer interaction and learning in cooperative small groups. *Jnl. of Educational Psychology, 74*(5) p.642-655

WEBB NM (1982b). Student interaction and learning in small groups. *Review of Educational Research, 52*(3) p.421-445

WEBB NM (1989). Peer interaction and learning in small groups. *International Jnl. of Educational Research, 13*(1) p.21-39

WEST LHT (1980). Towards descriptions of the cognitive structures of science students. In Archenhold, W. F., Driver, R. H., Orton, A., & Wood-Robinson, C. (Eds.). *Cognitive Development Research in Science and Mathematics.* Leeds: Leeds University.

WEINSHEIMER JC (1985). *Gadamer's Hermeneutics: a reading of Truth and Method.* New Haven: Yale University Press

WHEATLEY GH (1991). Constructivist perspectives on science and mathematics learning. *Science Education, 75*(1) p.9-21

WHITE RT (1988). *Learning Science.* Oxford: Blackwell

WHITE RT & GUNSTONE RF (1992). *Probing Understanding.* London: Falmer

WIERSMA W (1969). *Research Methods in Education: an introduction.* Philadelphia: JB Lippincott

WILLES MJ (1983). *Children into Pupils.* London: RKP

WINOGRAD T (1985). What does it mean to understand language? In Aitkenhead AM & Slack JM (Eds.). *Issues in Cognitive Modeling.* London: Lawrence Erlbaum

WITTGENSTEIN L (1922). *Tractatus Logico-Philosophicus.* Reproduced in Kenny A (Ed.) (1994). *The Wittgenstein Reader.* Oxford: Blackwell

WITTGENSTEIN L (1967). *Philosophical Investigations* (3rd edn), translated by GEM Anscombe. Oxford: Blackwell

APPENDIX

A. Teaching Concept Mapping

The following procedure has been developed to teach concept mapping to primary age children, and can be used with large groups, up to a whole class.

What you will need:

A sheet of paper and pencil for each child.

Make enough copies of the following "weather concepts" list so that each child in the group can have a strip of paper with the set of labels. (The children can either glue their labels in place on their sheet of paper, or they can copy the words once they have decided on a suitable arrangement.)

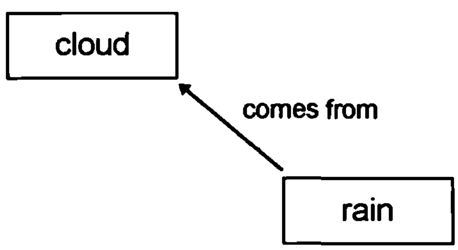
cloud rain puddles sunshine wind cold hot weather

For the second map, choose a set of around six to eight related concepts, ideally ones that the children have worked with recently, and prepare a similar set of strips. The following list of concepts works well with upper primary age children:

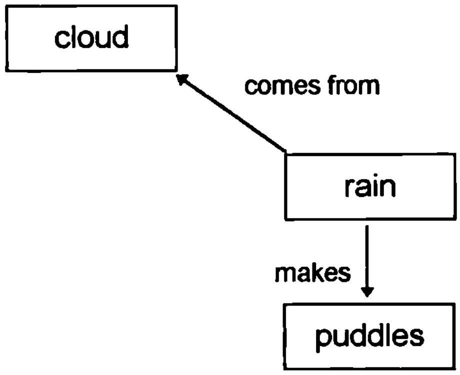
steam solid water substance liquid ice gas snow

What to Do:

1. Start with the map of weather concepts. Have the children tear or cut off the first word, "cloud", and the second word, "rain". Ask how the two words are related. Write up the word cloud and the word rain, and draw an arrow from rain to cloud. Write "comes from" along the arrow. Next have the children stick down their own labels near the middle of their paper, and join them in a similar way. Check that the arrows are drawn in the correct direction:

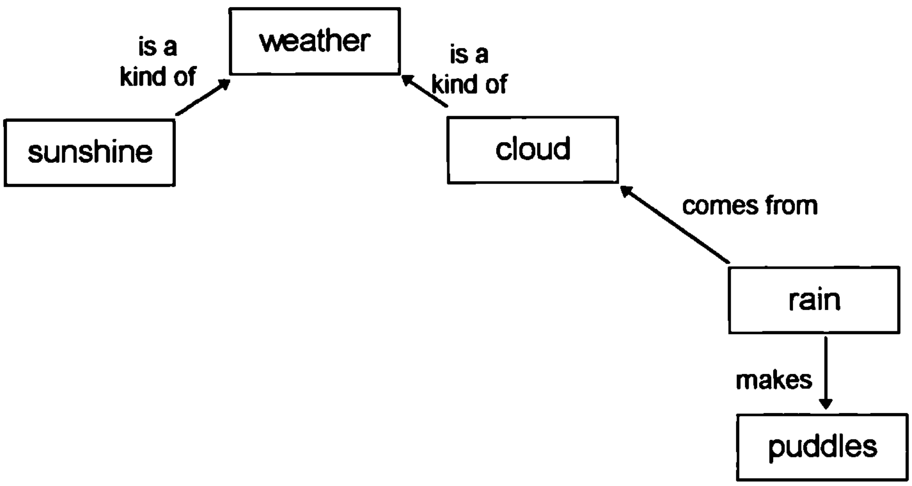


2. Now introduce the next word, "puddles". Ask how this joins to the previous words. Draw the connection, label it and have the children follow with their own maps:

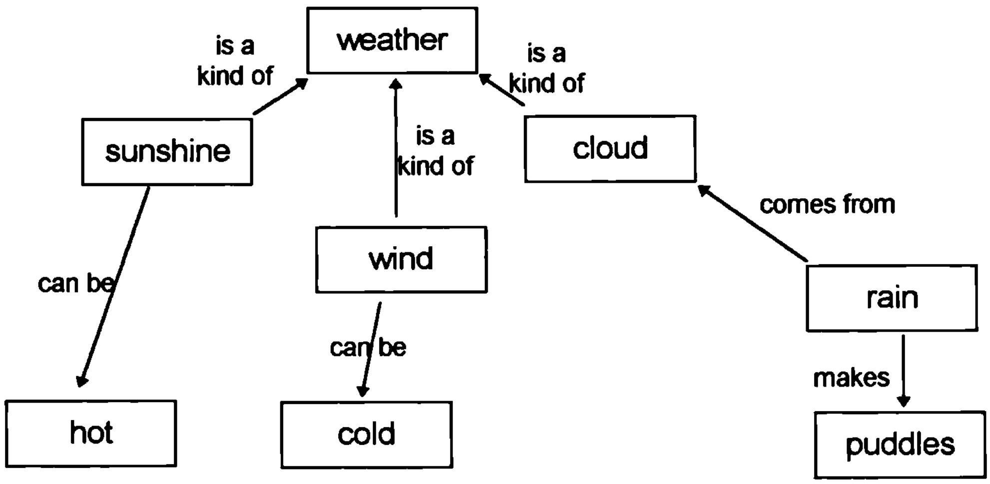


3. Ask where "sunshine" could be linked in. The children may have some suggestions, but say that they don't need to decide yet. Have them cut off the remaining words, and arrange them on their paper how they think they link in best.

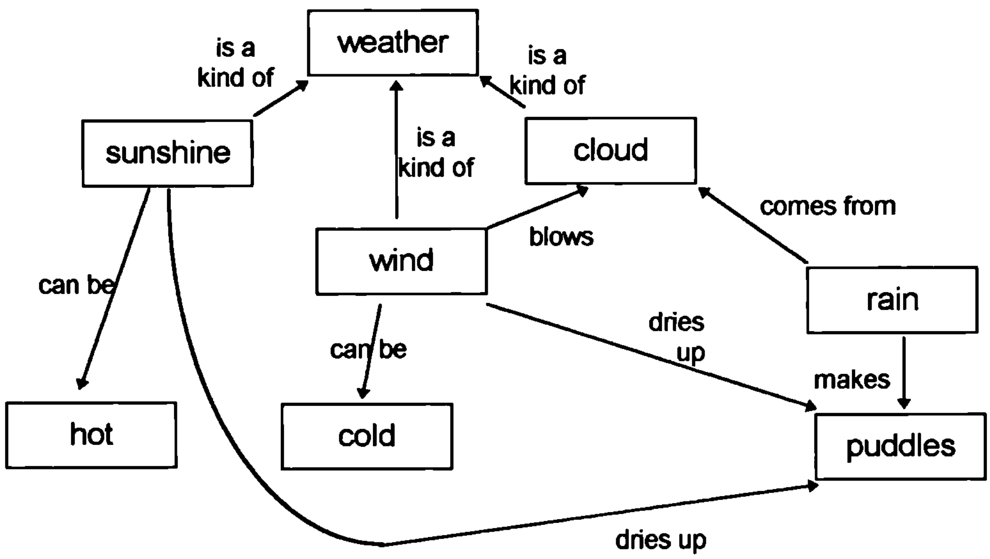
4. Point out that the word "weather" is an important word in this list, because it is the main idea behind all the other words. Say that because of this, it should go at the top of the map. Write up the word weather, and join it up:



5. Introduce the other words in turn, and ask where they link in. Have the children arrange them on paper. Write up the following, but stress that there are other arrangements that are equally good.



6. Show that, now all the words are in place, there are more links that can be made:



7. Encourage the children to think of further links that could be made.

8. Ask what further words could be added to the map. Depending on children's level of understanding, words such as "wet", "dry" or "vapour" could be incorporated.

9. When this map is completed, work may start on the second concept map. Give the children a copy of the prompt card below to guide them through the steps, reminding them as appropriate of corresponding steps in the first map. Start them off by asking them to decide which of the words should be near the top of the map. The only procedural difference is that children begin by deciding which words go near the top of the map, and which ones near the bottom.

10. Show some completed examples, emphasizing that there is no single correct solution. Discuss improvements that could be made to show more relationships on the maps. Discuss the children's choice of hierarchy.

Prompt card:

How to make a concept map

1. Look at the list of words
2. Decide which words should go at the top
3. Decide which words should go near the bottom
4. Put the top two or three words on your sheet of paper
5. Draw arrows to join up the words that go together
6. Write along the arrows how the words go together
7. Add the other words a few at a time, and join them up
8. You may add new words of your own to the map

Remember: there are many ways to make a map

B. Proposition Generation Task

The Acid Test

Use these terms in your answers:

---

acid alkali base hydrogen indicator neutral pH value salt substance water

---

1. Write as many sentences as you can about **acids,** using the words above

2. Write as many sentences as you can about **alkalis,** using the words above

3. Write as many sentences as you can, using the words above, to describe **what happens when an acid and an alkali are mixed**

C. Word Association Test

Instructions

Please write your name on the front of your booklet at the top.

Imagine that you have four test tubes.

One contains white vinegar, one contains washing soda, one contains salt, and one contains sulphuric acid.

They are all mixed with a lot of water, and they look the same.

Think of a safe way to find out which is which.

Now write one sentence on the front of your booklet that tells me what you would do.

If you managed to think of an idea, you had to try and remember what you knew about those substances.

When you hear or see a word, it often makes you think of other words. I'm going to read out some words used in science. You will find all the words printed in your booklet. As I read out each word, say the word to yourself, and write down the first word it makes you think of, on the first dotted line. For example [write on board], if I said the word "electricity", you might write down "light", because it made you think about an electric light.

Then say the word to yourself again, and write down the next word it makes you think of, on the next line. Keep saying the same word to yourself and writing down what it makes you think of until I tell you to go on to the next page. There aren't any right answers; I only want to know what you think of when you hear these words. Don't worry about spelling mistakes, just write down your ideas as quickly as you can; but please don't write the same answer twice for the same word.

Test Pages

acid ____________	neutral ____________
acid ____________	neutral ____________
acid ____________	neutral ____________
acid ____________	neutral ____________
acid ____________	neutral ____________
acid ____________	neutral ____________
acid ____________	neutral ____________
acid ____________	neutral ____________
acid ____________	neutral ____________
acid ____________	neutral ____________

hydrogen ____________	salt ____________
hydrogen ____________	salt ____________
hydrogen ____________	salt ____________
hydrogen ____________	salt ____________
hydrogen ____________	salt ____________
hydrogen ____________	salt ____________
hydrogen ____________	salt ____________
hydrogen ____________	salt ____________
hydrogen ____________	salt ____________
hydrogen ____________	salt ____________

alkali ________	indicator __________
alkali ________	indicator __________
alkali ________	indicator __________
alkali ________	indicator __________
alkali ________	indicator __________
alkali ________	indicator __________
alkali ________	indicator __________
alkali ________	indicator __________
alkali ________	indicator __________
alkali ________	indicator __________

water _________	substance __________
water _________	substance __________
water _________	substance __________
water _________	substance __________
water _________	substance __________
water _________	substance __________
water _________	substance __________
water _________	substance __________
water _________	substance __________
water _________	substance __________

base __________	pH value __________
base __________	pH value __________
base __________	pH value __________
base __________	pH value __________
base __________	pH value __________
base __________	pH value __________
base __________	pH value __________
base __________	pH value __________
base __________	pH value __________
base __________	pH value __________

D. Standardized Change-Score Analysis

1 Test of growth model

The appropriateness of the fan spread model may be evaluated using the following test.

(i) Calculate

$$t = \frac{(s^2_{y_w} - s^2_{x_w}).\sqrt{(N-4)}}{2.s_{x_w}.s_{y_w}.\sqrt{(1 - r^2_{xy_w})}}$$

Where:

s_{x_w}, s_{y_w} are the pooled within-group standard deviations of the pre- and posttest scores

r_{xy_w} is the within group correlation between pretest and posttest scores

N is the total number of subjects

(ii) Compare t with the critical value for $\alpha = 0.30$, df = N - 4.

If $t <=$ critical value, reject fan spread hypothesis.

If the fan spread hypothesis is not rejected, then an SCSA approach is appropriate.

2 Analysis method

In SCSA, scores are transformed (standardized) to force posttest variance to equal pretest variance. Then an adjusted gain score is calculated. Conventional analysis of variance (ANOVA) is then applied to the gain scores, using group membership as the main effect.

(i) The adjusted gain score Y′ is calculated by:

$$Y' = Y - X.(s_{y_w}/s_{x_w})$$

where:

Y is the posttest score

X is the pretest score

s_{y_w} is the pooled within-group standard deviation on the posttest

s_{x_w} is the pooled within-group standard deviation on the pretest

(ii) These adjusted gain scores are then analysed by carrying out a conventional ANOVA, but with degrees of freedom = 1, N-3.

E. Example Transcript

The following transcript is of group 7's pre-topic concept mapping session on the "Earth in space".

The transcript is broken into individual discourse moves, which are numbered down the left edge of the figure. The speaker is referred to here as the "actor", in keeping with the view of language adopted in Chapter 8. Actors are identified by a code, P1 to Pn for the pupils in the group, CA for the classroom assistant, RE for the researcher and TE for the teacher.

The wording is as close to that spoken by the children as possible, given that spoken English can differ somewhat in the words used from written English. In the interest of clarity, minimal transcription conventions have been introduced. Short pauses and hesitations are indicated by commas, in a similar way to how commas are used in writing. Longer pauses are indicated by an ellipsis. Where speech overlaps, this is indicated by underlining the overlapping segments. To preserve anonymity, names are replaced by asterisks.

The category into which each move is classified is shown next to the move. Alongside the category of move is shown the structure of the ideational exchanges in the discourse. Each exchange is depicted by a grey bar. The boundaries are marked by a black line, and each exchange is numbered sequentially for ease of reference. The paler parts of the bar indicate where the ideational exchange in question is no longer the focus of the discussion, and enable the reader to trace the continuity in the development of ideas across segments of the discourse during which they are dormant.

The final column in the figure provides a commentary, which interprets the progress of the discussion. This is intended to explain how the classification was made for each move. Often, the classification draws on paralinguistic cues such as intonation to aid this interpretation, and these are indicated where appropriate. References to the construct terms being mapped are signified in the commentary by <angle brackets>.

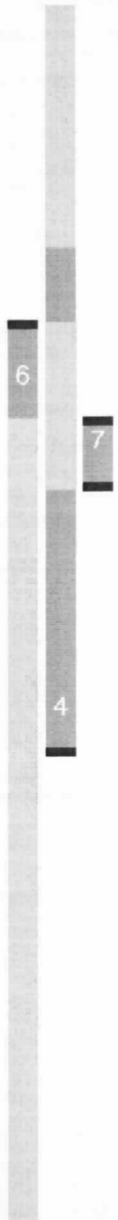
Following the transcript the concept map produced by this group is reproduced.

Move	Actor	Wording	Category	Exchange structure	Commentary
1.	P1	Shall we put universe at the top then?	Introducing	1	P1 introduces <universe> as the main idea for the map.
2.	P2	Yep	Supporting		This is generally welcomed, and P3 expresses on behalf of the group the justification for this. Intersubjectivity appears to have been achieved in this matter.
3.	P3	Yeah cos that's the that's like the one that's got all the stuff in	Elaborating		
4.	P1	Yeah exactly	Supporting		P1 confirms that intersubjectivity has been achieved.
5.	P4	Haven't got universe	Contextual		A contextual exchange is embedded within the ideational as there is some discussion about whether <universe> is one of the construct terms on their list. This is settled, and subsequently P3 says the word, sotto voce, whilst writing it onto the map.
6.	P3	Yes we have	Contextual		
7.	P3	Universe	Contextual		
8.	P4	Right, you got universe as the key word	Supporting		P4 reiterates support for <universe> as the 'key' idea, establishing this as the beginning of a hierarchy and orienting the discussion towards the next idea.
9.	CA	OK?	Contextual		The classroom assistant checks the group are alright.
10.	P2	[?]	Unclear		P2 seems to make a joke.
11.	P1	Right then ... got universe, so, and we've got the other words are stars, planets, Earth, moon, satellite, sun	Contextual		P1 reviews the construct terms they have been given.
12.	P2	Stars is like the second so	Introducing	2	P2 proposes <stars> as the next layer in the hierarchy, and immmediately P1 calls for justification for this.
13.	P1	Why?	Eliciting elaboration		

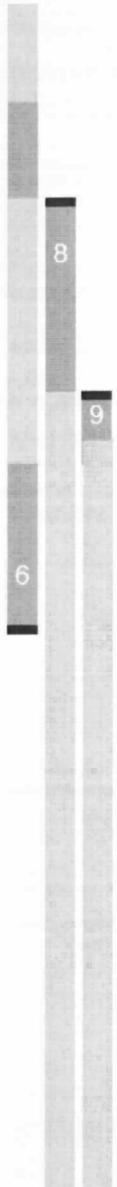
14.	P4	Why?	Supporting		P4 supports P1's call for justification.
15.	P3	Cos like it's more than like ... planets, stars	Elaborating		P3 seems to say, though not very succinctly, that <stars> are more important than <planets> in the hierarchy.
16.	P1	Stars and the sun ... are both balls of fire	Elaborating		P1 takes a slightly different line, treating <stars> and <sun> as different entities, but similar in respect of their being 'balls of fire'.
17.	P3	Right universe, let's start off with universe	Supporting	1	P3 reiterates support for <universe> as the key idea, perhaps sensing that the discussion is losing direction.
18.	P4	Planets ... right	Introducing	3	P4 introduces a new term, <planets>, for consideration.
19.	P2	Planets?	Challenging		P2, sounding incredulous, challenges P4's move.
20.	P3	Right, I'm starting, who says we start off with star, stars are in the universe?	Elaborating		P3 tries to re-establish the connection between <stars> and <universe> by specifying the nature of the relationship.
21.	P1	Yeah	Supporting	2	P1 supports this move by P3.
22.	P4	Right, stars	Contextual		The group then make the connection on their copies of the map, rehearsing the wording as they proceed.
23.	P4	Stars	Contextual		
24.	P4	Right, we got stars	Contextual		
25.	P1	Right, everyone writ that? stars, are in, the	Contextual		
26.	P3	Universe	Contextual		
27.	P3	Planets, I reckon should be another arrow coming off the universe	Elaborating	3	With the previous exchange and the physical process of making the connection completed, P3 takes up the move by P4 to introduce <planets> by specifying where it should be connected and the direction of the relationship.

28.	P1	Yeah, and then from planets, Earth can come from planets, cos Earth is a planet	} Introducing	4	P1 then identifies the next link that would follow on, specifying the nature of the relationship. P3 follows this line of thought, and completes P1's move simultaneously with P1, indicating that intersubjectivity is achieved.
29.	P3	Is a planet			
30.	P4	So what's going on to planet?	Query		P4 checks on what has been decided.
31.	P2	We should put like, planets, and then do Earth attached to planets	Supporting		P2 takes up the idea, offering support.
32.	P1	Yeah, that's what we're going to do	Supporting		The connection is decided.
33.	P3	Right so we've done stars, universe ... planets ... we can put Earth	Contextual		They review progress so far, and make the connections on the map, including the one, taken as shared, but never explicitly stated, between <planets> and <universe>.
34.	P1	Planet star	Contextual		
35.	P1	Planets are in the universe	Contextual		
36.	P4	We've done ... now we're on planets	Contextual		
37.	P3	Oi, af, after star, shall we put s, the sun is a star?	Introducing	5	The next connection is identified, made explicit, and then quickly accepted by P4.
38.	P4	Yeah	Supporting		
39.	P1	Under universe, we'll put, the sun is a star	Contextual		This is spoken for the benefit of the tape recorder.
40.	P2	What's that?	Query		P2 checks on what has been decided.
41.	P1	The sun is a star	Answer		

42.	P4	Underneath star, we should put, sun, and you write, is a, as the, thing	Answer	
43.	P1	And planet	Unclassified	
44.	P2	The sun is a star	Contextual	
45.	P2	We could do er Earth	Supporting	P2 keeps the connection between <Earth> and <planet> alive.
46.	P3	The moon, the moon, the moon	Introducing	There is a short period when several of the group speak together, with two of them, P3 and P4, attempting to move on to a new topic of discussion. P4's offering comes to the fore, but is not taken up.
47.	P4	The moon is on the other side of the star	Introducing	
48.	P1	Right, the Earth is a planet	Supporting	Meanwhile, P1 and P2 are still trying to make sure the connection between <Earth> and <planet> is firmed up. P3 and P4 divert from their new directions to agree to this.
49.	P2	Earth attached to planet	Supporting	
50.	P3	Yeah	Supporting	
51.	P4	Yeah	Supporting	
52.	P3	That's sort of splitting up into two sort of roads there, isn't it?	Contextual	P3 remarks on the appearance of the emerging concept map.
53.	P1	Yeah	Contextual	
54.	P2	Yeah	Contextual	
55.	P1	Earth	Contextual	They review what they have achieved.
56.	P2	The Earth is a planet	Contextual	
57.	P4	Sun we've done	Contextual	



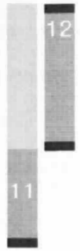
58.	P1	We've got two more, moon, and satellite	Contextual	
59.	P3	The moon is always rotating round the Earth	Elaborating	P3 returns to the attempt to connect in <moon>, this time elaborating on the nature of the relationship.
60.	P2	We might, universe, we might as well, we might as well put satellite attached to universe	Introducing	P2 attempts a different connection.
61.	P3	Yeah but what?	Eliciting elaboration	P3 is unable to see what the relationship could be, and asks for clarification.
62.	P2	No, Earth ... and the satellite	Introducing	P2, meanwhile, has seen a better link for <satellite>, and alerts the group.
63.	P4	Shall we put, moon, attached ... moon?	Supporting	P4 and P1 return to the connection between <moon> and <Earth>, and in the following moves this is transferred onto the map.
64.	P1	Underneath Earth, underneath Earth we'll put moon, rotates round the Earth	Supporting	
65.	P2	How do you spell rotates?	Contextual	
66.	P1	R, O, T, A, T, E, S, I think	Contextual	
67.	P3	Moon	Contextual	
68.	P2	What is it?	Query	
69.	P1	Rotates	Answer	
70.	P2	How do you spell it?	Contextual	
71.	P1	R, O, T, A ... T, E, S, I think	Contextual	
72.	P2	Yeah	Contextual	



73.	P3	So we've done Earth, moon, just need to do sat, s, we've done sun, satellite	Contextual		P3 recapitulates on where they have got to.
74.	P1	Satellite	Contextual		
75.	P2	Rotates around the Earth ... yeah?	Elaborating		P2 returns to the connection introduced earlier, elaborating on the nature of the relationship.
76.	P3	Cos a satellite can be a, big bowl of rock, or it could be a [?]	Elaborating		P3 recognizes that <satellite> does not just refer to artificial satellites. The implication seems to be that this is relevant to the connection being attempted.
77.	P4	Satellite is in the universe	Elaborating	8	P4 returns to the connection introduced by P2, specifying what the relationship could be.
78.	P2	The moon, the moon, the moon, em ... rotates around the Earth, I mean the satellite rotates around the Earth	Supporting		The others, however, pursue the connection between <satellite> and <Earth>.
79.	P3	Yeah but a satellite could be a big piece of rock, that can be a satellite, as well as a little bit of tin that gives out signals	Challenging		P3 gives renewed emphasis to the broader definition of <satellite>, seemingly implying that the connection may need to be reconsidered. Subsequently, though, P3 reconciles the perceived conflict and approves the connection.
80.	P3	Let's just put rotates around the Earth, cos they both do	Integrating		
81.	P2	Yeah	Supporting	9	
82.	P1	*[to neighbouring group]* Alright, J***			
83.	RE	How you getting on?	Contextual		The researcher arrives to see where the group have got to with their map.
84.	P4	OK	Contextual		
85.	P1	OK	Contextual		

86.	P2	Rotates around	Contextual		P2 rehearses the wording to be written on the map.
87.	P4	Now we underline ... now we can underline all of our words, so we know which ones are	Contextual		P4 is now thinking in terms of enhancing the appearance of the map.
88.	RE	Is this what you've just done?	Contextual		
89.	P3	Yeah	Contextual		
90.	P3	What we could do, we could put, look, a line across here, the sun is the Earth's star	Introducing	10	P3 draws attention to another potential connection.
91.	P4	Oh yeah	Supporting		P4 responds in a rather surprised tone of voice, approving P2's move.
92.	P4	So from sun to Earth you put a line, an arrow, saying, the sun is the Earth's star	Contextual		They then discuss how physically to make the connection. P1 is experiencing difficulty fitting this into the configuration of the map, and the others offer solutions.
93.	P3	The Earth's star	Contextual		
94.	P1	From sun to Earth, did you say?	Contextual		
95.	P1	I got a bit of trouble here now	Contextual		
96.	P4	Now you should put sun is the star, for the Earth	Supporting	10	
97.	P3	Just put it here	Contextual		
98.	P1	Look	Contextual		
99.	P1	Do I put, sun, how can you go to Earth?	Contextual		
100.	P4	Just put like that, and like that	Contextual		

101.	P3	Oi, connect it up with moon and planets and all. Cos all planets have moons	Introducing	11	P3 introduces another connection for <moon>, quickly backing it up with a justification.
102.	P4	[?] what is it?	Query		
103.	P1	Say that again?	Query		The others seem not to have grasped what P3 has said, and they get P3 to repeat it. P4 seems still to be thinking about a previous connection. However P1 continues to query what was said.
104.	P4	The sun, is, the	Answer		
105.	P1	Say that again about the moon and planets?	Query		
106.	P3	Cos all, most of the planets have a moon, so if we connect	Elaborating		P3 responds to the query, but modifying what had originally been proposed.
107.	P2	The moon attach, to planets?	Query		P2 now queries what connection is being suggetsed
108.	P3	Yeah	Answer		
109.	P2	We've already attached moon to something	Contextual		The reason for P2's query seems to be a concern not to make too many connections to one term, but the others assure P2 that this is quite acceptable.
110.	P3	I know, but you can attach it twice	Contextual		
111.	P1	Yeah you can attach it again	Contextual		
112.	P1	What shall I write?	Eliciting elaboration		P1 has a more substantive question regarding the nature of the relationship intended, and this is backed up by P4.
113.	P2	Moon to planets	Supporting		
114.	P4	That's what I mean ... what can you write?	Supporting		
115.	P3	Planet, planets have a moon	Supporting		P3 reiterates the relationship proposed.

116.	P2	The moon is a planet ... will do just for that	Introducing		Just as P3 repeats the idea, P2 proposes an alternative idea: a misconception. Subsequently, P2 repeats this idea.
117.	P2	The moon is a planet I'm writing	Supporting		
118.	P3	The planets have a moon, that's what I've writ	Supporting		The misconception is not challenged directly. P3 simply reaffirms the original proposal.
119.	P3	But it's all simple on that side, but when you get to here there's loads of it	Contextual		P3 observes that the emerging map is asymmetric.
120.	P1	Yeah	Contextual		
121.	P3	*[to neighbouring group]* J***, you're not supposed to do that			
122.	TE	*[speaks to neighbouring group]*			
123.	P4	Miss, what do you mean by, add other words from the list?	Contextual		P4 queries with the teacher the meaning of an instruction given to them with their list of construct terms. This is cleared up, and they then turn to transcribing the rough versions of the map into the final draft.
124.	TE	Sorry?	Contextual		
125.	P4	What do you mean by add other words from the list?	Contextual		
126.	TE	Well you use all those words that are on there	Contextual		
127.	P1	Well we've used them	Contextual		
128.	P3	We've got to write it up on the big one	Contextual		
129.	TE	Yeah, you need to write it up	Contextual		

Concept Map for Group 7 (pre-topic)